Uniwersytet Warszawski
Wydział Biologii

Sylwia Szczepaniak
216676

The contribution of *Saccharomyces cerevisiae* cap binding complex to small nuclear and small nucleolar RNA processing

Rozprawa doktorska
w zakresie nauk biologicznych
dyscyplinie biologii

Praca doktorska wykonana pod kierunkiem Prof. dr hab. Joanny Kufel (Instytut Genetyki i Biotechnologii, Wydział Biologii UW) oraz Prof. dr hab. Edwarda Darżynkiewicza (Zakład Biofizyki IFD, Wydział Fizyki UW) w ramach Międzywydziałowych Interdyscyplinarnych Studiów Doktoranckich

Warszawa, grudzień 2015

Oświadczenie kierującego pracą

Oświadczam, że niniejsza praca została przygotowana pod moim kierunkiem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie stopnia doktora nauk biologicznych w zakresie biologii.

Data Podpis kierującego pracą

Oświadczenie autora pracy

Świadom odpowiedzialności prawnej oświadczam, że niniejsza rozprawa doktorska została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Oświadczam również, że przedstawiona praca nie była wcześniej przedmiotem procedur związanych z uzyskaniem stopnia doktora w innej jednostce.

Oświadczam ponadto, że niniejsza wersja pracy jest identyczna z załączoną wersją elektroniczną.

Data Podpis autora pracy

Słowa kluczowe

RNA maturation, 3’ end processing, m^7G cap, TMG cap, CBC, NNS, Rnt1, snRNA, snoRNA.

Tytuł pracy w języku polskim

Udział kompleksu wiążącego kap w dojrzewaniu małych jądrowych i małych jąderkowych RNA u drożdży *Saccharomyces cerevisiae*”.

Abstract

Small nuclear and nucleolar RNAs are transcribed by RNA polymerase II as precursors, whose 5' ends either undergo combined endo- and exonucleolytic processing by Rnt1 and Rat1, respectively, or remain unchanged. In turn, their extended 3' ends are trimmed by the exosome/Rrp6, following cleavage by Rnt1 or transcription termination by the ncRNA-specific transcription termination Nrd1/Nab3/Sen1 (NNS) complex. In this work showed that the interaction between Nrd1/Nab3 and the cap binding complex (CBC) bound to the m^7G cap of sn/snoRNA precursors is a key element of the mechanism linking their 5'- and 3' end processing. I demonstrated that CBC binds to the m^7G cap at the early step of transcription but remains associated with snoRNAs and snRNA genes, also for snoRNAs processed by Rnt1, until the termination and is dependent on RNA. Moreover, in the absence of the CBC complex Nrd1 recruitment to sn/snoRNA genes is decreased. I also showed that co-transcriptional recruitment of Rnt1 to snoRNA genes processed at their 5' end occurs at late stages of transcription. Surprisingly, in the absence of Rnt1, 5' end unprocessed box C/D pre-snoRNAs do not possess mature 3' ends but carry oligoadenylated extensions, which indicates a defect in their 3' maturation. These data strongly suggest that processing of both snoRNA termini is coupled, probably via the CBC-NNS interaction, that prevents premature 5' end processing by Rnt1 and delays it until transcription termination. This link probably contributes to the error-free and timely synthesis of mature snoRNA molecules.

Streszczenie

Małe jądrowe i jąderkowe RNA (sn/snoRNA) są transkrybowane przez polimerazę RNA II (Pol II) jako prekursory, których końce 5' ulegają obróbce endo- i egzonukleolitycznej z udziałem nukleaz Rnt1 i Rat1, lub pozostają niezmienione. Z kolei wydłużone końce 3' powstałe w wyniku cięcia katalizowanego przez Rnt1 lub terminacji transkrypcji zależnej od specyficznego dla niekodujących RNA (ncRNA) kompleksu Nrd1/Nab3/Sen1 (NNS) są skracane z udziałem jądrowego egzosomu/Rrp6. W pracy wykazano, że oddziaływanie NNS z kompleksem wiążącym m^7G kap (CBC) obecny na końcu 5' prekursorów sn/snoRNA jest kluczowym elementem mechanizmu łączącego dojrzewanie końców 5' i 3'. CBC wiąże kap na wczesnych etapach transkrypcji genów sn/snoRNA i pozostaje z nim związany aż do terminacji transkrypcji, również w przypadku snoRNA, których koniec 5' jest cięty przez Rnt1. Ponadto, w nieobecności CBC, rekrutacja Nrd1 do genów sn/snoRNA jest obniżona lub opóźniona. Pokazano również, że asocjacja Rnt1 do genów snoRNA ciętych na końcu 5' zachodzi ko-transkrypcyjnie ale na późnych etapach transkrypcji. Co ciekawe, w nieobecności Rnt1, prekursory snoRNA klasy box C/D, które zostały przecięte na końcu 5', mają niedojrzały, wydłużony i oligoadenylowany koniec 3', co sugeruje defekt jego obróbki. Dane te silnie wskazują na związek między dojrzewaniem obu końców sn/snoRNA, prawdopodobnie poprzez oddziaływanie CBC-NNS, które zapobiega przedwczesnemu cięciu Rnt1, opóźniając je aż do momentu terminacji transkrypcji. Opisany mechanizm jest prawdopodobnie istotnym elementem kontroli jakości i synchronizacji dojrzewania pre-sn/snoRNA.

Acknowledgements

I would like to express my deepest gratitude to my PhD Thesis Supervisor Professor Joanna Kufel for her guidance, wisdom and support not only professional but also in countless personal situations.

I would like to thank my Supervisor Professor Edward Darżynkiewicz for his attention and insight during the research and preparation of this thesis.

I would like to thank all the members of Professor Joanna Kufel's Laboratory for all their help, support and wonderful, warm atmosphere in the lab. Particularly I would like to acknowledge MSc Dorota Adamska and BSc Żaneta Matuszek for their contribution in this project.

I would also to thank Professor Andrzej Dziembowski and members of his laboratory for their expertise and technical help. Especially I would like to acknowledge MSc Katarzyna Kowalska and MSc Krystian Stoduś for their contribution in plasmid and protein purification.

This work was supported by the Polish Ministry of Science and Higher Education (N N301 065740), National Science Centre (UMO-2011/01/N/NZ1/04344), Polish-Swiss Research Programme (PSPB-183/2010; Functional diversity of non-coding RNAs) and the EU through the European Social Fund (UDA-POKL.04.01.01-00-072/09).

Experiments were carried out with the use of CePT infrastructure financed by the European Union- the European Regional Development Fund (Innovative economy 2007-13, Agreement POIG.02.02.00-14-024/08-00).

During my PhD I received the EMBO Short-Term Fellowship, which allowed me to work in Shona Murphy's and Nicholas Proudfoot's laboratories (Sir William Dunn School of Pathology, University of Oxford, UK).

Table of contents

Frequently used abbreviations

APT – Associated with Pta1, subunit of the cleavage and polyadenylation factor (CPF)

asRNA – antisense RNA

CBC – Cap Binding Complex

CBCA – CBC-ARS2 complex

CBCN – CBC-NEXT complex

ChIP – Chromatin ImmunoPrecipitation

CF – Cleavage Factor

CP – Cleavage and Polyadenylation

CPF – Cleavage and Polyadenylation Factor

cRT-PCR – circularized Reverse Transcription PCR

CTD – C-Terminal Domain of polymerase II

CUT – Cryptic Unstable Transcript

EC – Elongation Complex

Exo-9 – Exosome 9-subunit complex (without Dis3)

Exo-10 – Exosome 10-subunit complex (Exo-9 plus Dis3)

Exo-11 – Exosome 10-subunit complex (Exo-10 plus Rrp6)

ncRNA – noncoding RNA

NEXT – Nuclear Exosome Targeting

NNS – Nrd1/Nab3/Sen1 complex

NUT – Nrd1-Unterminated Transcripts

m^7G – 7-MethylGuanosine cap

mRNA – messenger RNA

PAS – PolyAdenylation Site

Pol I – Polymerase RNA I

Pol II – Polymerase RNA II

Pol III – Polymerase RNA III

P-TEFb – positive Transcription Elongation Factor b

RNP – RiboNucleoprotein Particle

rRNA – ribosomal RNA

Ser2-P – Phospho-Serine 2 (of CTD)

Ser5-P – Phospho-Serine 5 (of CTD)

Ser7-P – Phospho-Serine 7 (of CTD)

snoRNA – small nucleolar RNA

snRNA – small nuclear RNA

SUT – Stable Unannotated Transcripts

TMG – $m^{7,2,2}$-TriMethylGuanosine cap

TRAMP – Trf4/Air2/Mtr4p Polyadenylation complex

tRNA – transfer RNA

TSS – Transcription Start Site

1. INTRODUCTION

According to recent studies up to 90% of eukaryotic genome is transcribed, but protein coding genes comprise only a small proportion (2-3%) (1). The majority of the transcriptome is untranslated and consists of noncoding RNAs (ncRNAs), which are no longer considered as "transcriptional noise", since there is accumulating evidence that many of them have important biological functions. The best characterized examples of ncRNAs are small nuclear and small nucleolar RNAs (sn/snoRNAs) that, like protein coding mRNAs, are synthesized by RNA polymerase II (Pol II). Although our knowledge of Pol II transcription and co-transcriptional processing is extensive for protein coding genes, much less is known about the biogenesis of Pol II transcribed ncRNAs. The subject of this PhD thesis is the maturation of the 3' and 5' ends of sn/snoRNAs in yeast *Saccharomyces cerevisiae*.

1.1 Coordination of co-transcriptional events

All eukaryotic RNAs are synthesized as precursors and their processing provides an important level of regulation of their expression. Most processing factors as well as components of mature ribonucleoprotein particles (RNPs) are recruited co-transcriptionally in a tightly regulated manner. Even assembly of factors that act at the latest stages of transcription (e.g. 3' end processing machinery) or post-transcriptionally often occurs much earlier during transcription. The transcription cycle and transcription related events are organized by post-translational modifications of the C-terminal domain (CTD) of Pol II (2–5). Additional mechanisms such as transcription elongation rate tuning, Pol II pausing and the complex network of interactions between transcription and processing factors also contribute to the synchronisation of these processes (6).

1.1.1 Pol II C-terminal domain

The C-terminal domain of the largest Pol II subunit Rpb1 is a distinctive feature of this polymerase (2–5). It is dispensable for Pol II activity *in vitro*, and its partial deletion can be tolerated, although deletion of the entire CTD in mice, Drosophila or yeast is lethal (2). CTD orchestrates Pol II transcription and transcription related events serving as a scaffold for signalling between the transcription machinery and multiple auxiliary factors. It was shown to be involved in regulating transcriptional rates, chromatin modifications, RNA co-transcriptional processing, including capping, splicing and 3' end formation, transcription termination and finally RNA export, as well as in DNA replication and maintenance of

genomic stability. CTD is composed of multiple tandemly repeated heptapeptides with the consensus sequence $Y_1S_2P_3T_4S_5P_6S_7$. The number of repeats varies among different species and correlates with the organism developmental level (from 26 repeats in yeast to 52 in humans) and deviations from the consensus (2, 5). Another important feature of the CTD, at least in humans, is the presence of the Rpb1-stabilizing 10-amino- acid non-consensus motif following heptad 52 (3). Due to such a composition, CTD possesses a huge structural and functional plasticity. This long and flexible tail can reach different elements of the actively transcribed chromatin (3). The reach of the CTD as well as its affinity to particular factors is further regulated by post-translational modifications. The phosphorylated tail could be extended even six times, being four times longer than the Pol II surface diameter (7). In theory, each residue can be modified, with 5 of them being phosphorylated. Additionally either of the two prolines can exist in a *cis* or *trans* conformation. This so-called "CTD code" was extensively studied over the last decade (2–5).

The best characterized CTD modifications are these of the three serine residues: Serine 5 (Ser5-P), Serine 2 (Ser2-P) and Serine 7 (Ser7-P) (2–5). The pattern of these modifications dynamically changes during the transcription cycle, which governs the association of stage-specific factors (Fig. 1A and B). High-throughput studies, mainly using chromatin immunoprecipitation (ChIP) techniques, established characteristic Ser5-P, Ser2-P and Ser7-P profiles along genes and intragenic regions. Generally, Ser5 phosphorylation seems to be higher at the beginning of the transcription cycle and becomes reduced towards the end, whereas Ser2 phosphorylation increases towards the end of the transcription unit (Fig. 1A) (8–10). This pattern is important for synchronizing transcription of all Pol II transcribed genes, whereas the far less understood Ser7 phosphorylation seems to play a more specific function. However, it is noteworthy, that recent studies, using novel native sequencing technology for elongating transcripts, showed significant differences in CTD phosphorylation profiles across mammalian protein coding genes. Particularly most Ser5-P was detected not over the transcription start site (TSS) region of the gene, but in the gene bodies, where it is associated with actively spliced exons (11). There are several other differences between yeast and higher eukaryotes, and the description below is focused on the situation in *S. cerevisiae*.

At the beginning of the transcription cycle Ser5 is phosphorylated by Kin28 (CDK7 in higher eukaryotes) (Fig. 1A). *In vitro*, Ser5 has been shown to be phosphorylated also by mammalian CDK8, a subunit of the generic transcription regulator complex, that can phosphorylate Ser2 (12). However, yeast CDK8 homologue, Srb10, does not contribute to

CTD phosphorylation *in vivo*. During transcription elongation Ser5-P is removed by Ssu72 and Rtr1 phosphatases (Fig. 1A) (7, 13).

Rtr1 associates with genes at the point where Ser5-P levels drop and Ser2-P starts to increase, and its inactivation leads to elevated Ser5-P across the coding region. Its human homologue, RNA Pol II associated protein 2 (RPAP2), is essential for efficient transcription and 3' end processing of snRNA transcripts. In turn, Ssu72 which is the component of the APT (associated with Pta1) subunit of the cleavage and polyadenylation factor (CPF) is stimulated by Pta1, associates with the elongating complex (EC) at both the promoter and more robustly at the 3' end region, and its depletion results in increased Ser5-P towards the 3' end (14–16). The main role of the Ser5-P modification, demonstrated specifically in *Schizosaccharomyces pombe*, is the recruitment of the capping enzyme (Fig. 1B) (see 1.2.1) (17). However, also other important factors require the Ser5-P mark for association with the transcription machinery, including the Set1-COMPASS histone methyltransferase, Rtr1, the histone deacetylase Rpd3S and the termination factor Nrd1 (Fig. 1B) (see 1.3.2) (2).

Also Ser7 phosphorylation takes place early during the transcription cycle, and is performed by Kin28, but also by Ser2-P kinase Bur1 during transcription elongation (2, 9). Accordingly, Ser7-P profiles generally resemble those of Ser5-P at the beginning of yeast genes, but in contrast to Ser5-P, Ser7-P level remain high until transcription termination (Fig. 1A) (9, 10). Ser7-P profile is similar on protein coding and noncoding Pol II transcribed genes (Fig. 1A) (8, 9). Ser7-P mark is removed by the Ser5-P Ssu72 phosphatase, but not by Rtr1, immediately after cleavage and polyadenylation, contributing to the reconstitution of the hypo-phosphorylated state of Pol II (Fig. 1A) (14, 15, 18, 19). Ser7-P may be important for elongation, as high levels of Ser7-P are detected on highly transcribed genes (8, 9). Genome-wide studies in yeast showed that Ser7-P profile often colocalize with Nrd1 at the 3' ends of cryptic unstable RNAs (CUTs) and stable unannotated transcripts (SUTs) and within introns encoding snoRNAs, however the meaning of this correlation is not clear at the moment (8). In humans, Ser7-P is required for the expression of snRNA genes by facilitating the interaction with the snRNA-specific 3' end formation complex, the Integrator (20). Ser7-P also facilitates the recruitment of the CTD Ser5 phosphatase RPAP2, associated with a catalytically inactive subcomplex of the Integrator, to snRNA promoter regions (21). RPAP2 dephosphorylates Ser5-P while Ser2 is phosphorylated, creating a 'double mark' composed of Ser7-P on one repeat and Ser2-P on the following repeat, which is then specifically recognized by Int11 (22).

Serine 2 phosphorylation varies the most between different transcription units, but in general it is significantly lower on noncoding genes (8, 9). Phosphorylation is carried out by two Ser2 kinases, Bur1 and Ctk1, that are essential for normal growth of yeast cells, with the latter playing a major role during Pol II elongation (Fig. 1A) (2, 23, 24). Besides CTD phosphorylation, the main role attributed to Bur1 is stimulation of elongation and suppression of aberrant initiation by phosphorylation of other transcription factors like Spt5, that contains a C-terminal repeat domain (CTR) similar to that in CTD (9, 24–26). Bur1 is recruited to the EC through Ser5-P CTD and Ser2 phosphorylation by Bur1 at promoter proximal regions further stimulates phosphorylation by Ctk1 (24). In higher eukaryotes, Cdk9, a subunit of the positive transcription elongation factor b (P-TEFb), phosphorylates both Spt5 and Ser2 (27). Additional CTD Ser2 kinases Cdk12 and Cdk13 have been identified recently in higher eukaryotes, with Cdk12 contributing the majority of Ser2-P during transcription elongation (28, 29). Also a bromodomain protein Brd4 is an atypical Ser2-P kinase, involved in P-TEFb recruitment, which phosphorylates Ser2 while P-TEFb is inactive (30). Yeast Ctk1 and Bur1 and human P-TEFb was recently shown to interact with cap binding complex (CBC), which stimulates their recruitment and thus Ser2 phosphorylation (see 1.2.2) (31–33). At the end of the transcription cycle Ser2-P mark is removed by Ser2 specific phosphatase Fcp1 (Fig. 1A) (7, 15), enabling the polymerase recycling. Although Fcp1 travels with Pol II along the gene and can dephosphorylate both Ser2-P and Ser5-P, it prefers Ser2-P (3). Ser2-P and Ser5-P are additionally dephosphorylated by phosphatase Cdc14 during mitosis to repress transcription (34).

The role of Ser2-P in Pol II elongation is not fully understood, but this modification was shown to be important for splicing, transcription termination and 3’ end formation (see 1.3) (2). Consistently, Ser2-P is placed downstream from the promoter and its level increases towards the 3’ end of genes (Fig. 1A) (8–10). CTD phosphorylated on both Ser2 and Ser5 recruits numerous transcription and RNA processing factors as well as histone modification enzymes, including splicing factors Prp40 and U2AF65, export factor Yra1, H3K36 methyltransferase Set2, the RNA binding factor Ssd1, the mitotic kinase Hrr25, and the RecQ5 genome stability helicase (Fig. 1B) (2). On the contrary, elongation-associated proteins, Npl3 and Spt6, as well as CTD-interaction domain-containing (CID) termination factors, Rtt103 (together with the 5’-3’ exonuclease Rat1) and Pcf11, associate with the CTD phosphorylated on Ser2 alone (Fig. 1B) (2, 3). Both termination factors, Rtt103 and Pcf11, bind cooperatively to neighbouring Ser2-P repeats, which ensures that they are efficiently

recruited only when many repeats have Ser2-P (35). Also Sen1 helicase, another Pol II termination factor, binds to the Ser2-P CTD (36, 37). Conversely, termination factors can also affect CTD phosphorylation, as was shown for exonuclease Rat1 (see 1.2.4.2) (38).

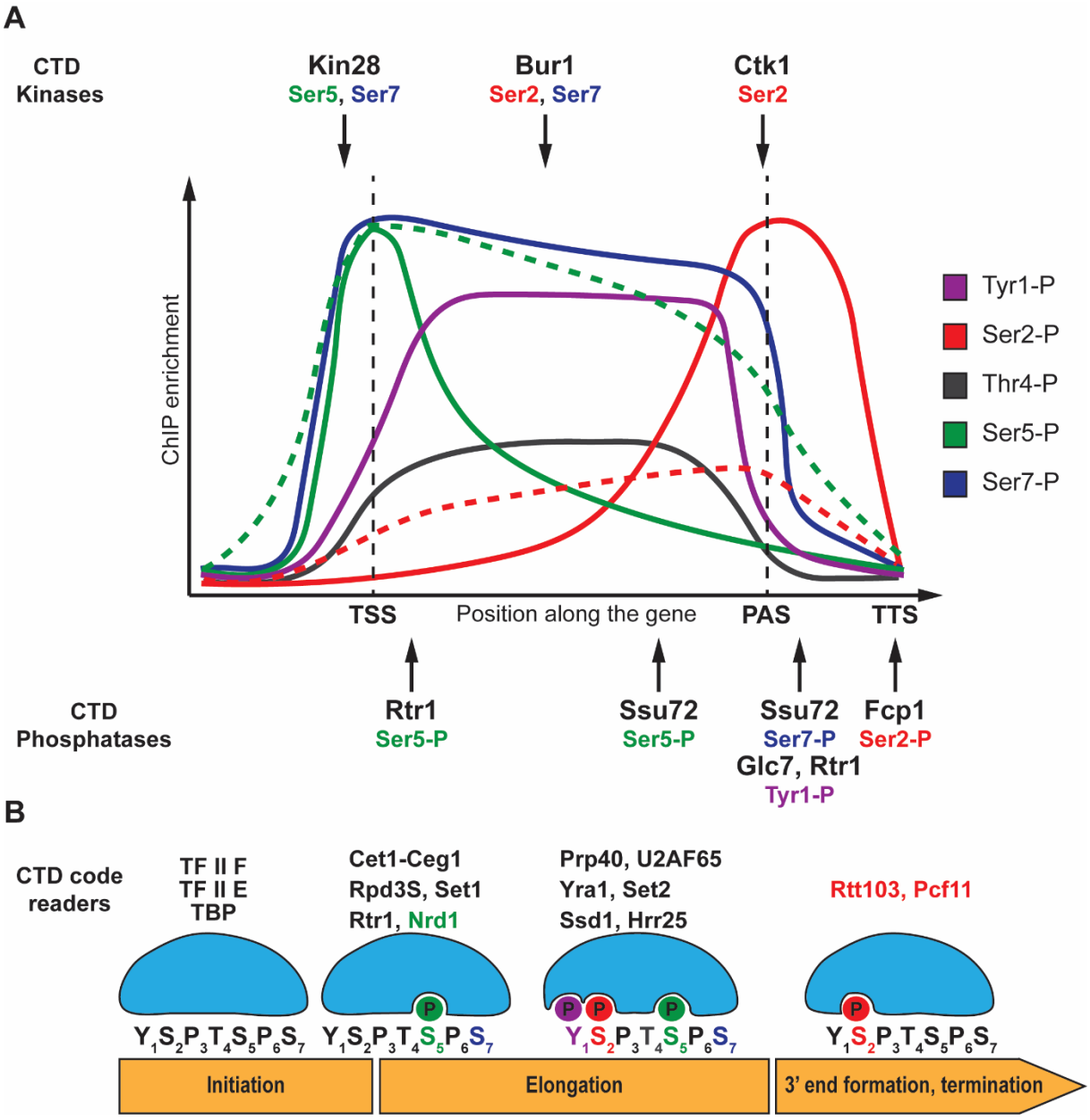

Figure 1. The carboxyl-terminal domain (CTD) phosphorylation pattern across yeast genes. *(**A**) Average profiles of Tyr1-P, Ser2-P, Thr4-P, Ser5-P and Ser7-P along typical protein coding genes (continuous lines) or snoRNA genes (dashed lines) visualized by ChIP experiments in S. cerevisiae according to* (39)*. Known kinases and phosphatases responsible for establishing and eradicating these marks are indicated above and below the diagram, respectively. Transcription start site (TSS), polyadenylation signal (PAS) and transcription termination site (TTS) are indicated on the x axis. (**B**) Selected factors binding specifically to the differentially-modified CTD (CTD code readers) at different stages of the transcription cycle. Termination factors are coloured in green and red. Description in the text.*

Although Ser2-P level at short sn/snoRNA genes is low, this modification plays important role in the Nrd1/Nab3/Sen1-dependent termination (see 1.3.2), as deletion of the CTK1 gene leads to extensive transcriptional readthrough of these genes (40). It can be partly explained

by the altered elongation kinetics caused by the disturbed recruitment of elongation factors and histone modifiers. However, according to recent studies proper Ser2-P level at the 3' end of noncoding transcription units is needed for the association of Pcf11 and Sen1 (Fig. 1B), which seems to be vital for the termination pathway mediated by the Nrd1/Nab3/Sen1 (NNS) complex (37, 41).

Much less is known about the significance of Thr4 phosphorylation that is limited to the gene body, where it may impair the recruitment of termination factors (Fig. 1A) (42). Lack of Thr4-P is not lethal in yeast but it is vital for transcription elongation in mammalian cells where its profiles overlap with Ser2-P with a slight shift towards the 3' end (43). Different results were obtained in chicken cells, where Thr4-P was shown to play a role in processing of non-polyadenylated histone mRNAs (44). Also the identity of the kinase that phosphorylates Thr4 is not entirely clear, either Ser2 kinase Cdk9 or Plk3 (Polo-like kinase 3) were implied to be involved, depending on the system, chicken or human, respectively (43, 44), whereas *S. cerevisiae* Thr4 kinase has not been identified yet.

Phosphorylation of Tyr1 was demonstrated in humans almost two decades ago, but only recently this modification and its function was analysed in yeast (42, 45). Tyr1-P is present at all active genes and its level correlates with the Ser2-P mark, increasing downstream of the TSS but dropping before reaching the polyadenylation signal (PAS) (Fig. 1A) (42). Tyr1-P was shown to inhibit the association of termination factors to the EC that probably serves to prevent premature termination. Therefore, termination factors peak upstream the highest Tyr1-P level (Nrd1) or further downstream, after Tyr1-P dephosphorylation (Pcf11, Rtt103) (Fig. 1B). Tyr1 kinase has not been identified yet, whereas APT subunit phosphatases Glc7 and Rtr1 were shown to remove the Tyr1-P mark (46, 47).

Proline isomerization, performed by proline isomerases Ess1 in yeast and Pin1 in mammals, contributes to the association of numerous factors and is particularly important for transcription termination in yeast. Pcf11 specifically recognizes the repeats with Ser2-P and prolines in the *trans* configuration (48), which is a dominant isomer in equilibrium (49). In contrast, Nrd1 and Ser5-P phosphatase Ssu72 bind CTD with Ser5-P and the downstream proline in the *cis* configuration (49, 50). It was shown that proline isomerization by Ess1 activates Ssu72, facilitating Ser5-P dephosphorylation (49, 51, 52). In turn, CTD phosphorylation state is important for the activity of Ess1 (53). In yeast, Ess1 mutation causes readthrough of many snoRNA genes, probably due to the disturbed recruitment of Nrd1 and Pcf11 (41, 51).

In higher eukaryotes not all CTD repeats have the consensus sequence. For instance, the mammalian CTD contains 31 non-consensus repeats, located mainly in the C-terminal region (2). They predominantly differ at position 7 where serine is often replaced by lysine, a potential substrate for acetylation, methylation, sumoylation and ubiquitination (2, 4). These repeats can be additionally, specifically modified and may contribute to specific CTD functions. For instance, the arginine in position 7 of repeat 31 (R1810) can be methylated during transcription, which probably serves to inhibit general transcription of sn/snoRNA genes (54).

Individual CTD modifications influence each other and these connections may enable the general transcription machinery to distinguish between "coding" and "noncoding" transcription mode. Several genome-wide Pol II ChIP studies in yeast demonstrated that usually longer genes have phosphorylation patterns more consistent with the general model, with Ser5-P high at the beginning and increasing Ser2-P towards the end (8–10, 14, 15). Shorter transcription units, most often noncoding, are characterized by higher levels of Ser5-P and lower levels of Ser2-P, probably due to the crossover of Ser5-P to Ser2-P occurring 450 nt downstream from the TSS (Fig. 1A). The meaning of this difference is underscored by different requirements for the recruitment of the coding versus noncoding termination factors (see 1.3). The association of the cleavage and polyadenylation (CP) machinery is Ser2-P dependent, whereas Nrd1 binds Ser5-P CTD (3). However, appropriate level of Ser2-P phosphorylation is also required for ncRNA transcription termination, as it contributes to Pcf11 and Sen1 recruitment (41). Additionally, it was proposed that a low level of Ser2-P and high level of Ser7-P serves as a CTD gene-type specific signal for noncoding transcription units (Fig. 1A) (2). The balance between Ser2-P and Ser7-P is also emphasized by the finding that mutating all Ser7 residues can bypass a gene-specific requirement for Lsk1-mediated CTD Ser2 phosphorylation in regulation of meiosis in *S. pombe* (17, 55). In turn, Ser5-P was shown to be important for subsequent phosphorylation of Ser2 (24, 56). Ser5-P is also tightly connected to Ser7-P, as these marks share a common kinase Kin28 and phosphatase Ssu72.

CTD is not required for Pol II catalytic activity, but its deletion is lethal and at least 50% of the natural numbers of heptads are necessary for cell viability (3). In yeast, eight heptads of total 26 repeats are sufficient for viability but 13 are required for wild-type growth (57). Although deletion of the CTD has no effects on transcription, it is required for efficient 5' capping, splicing, and 3' processing (58, 59). Also mutations of CTD residues that undergo modification, for instance substitution of all Tyr1 residues to phenylalanine (Phe) or of Ser2

or Ser5 to alanine (Ala) in *S. cerevisiae*, are lethal (17, 57). In turn, in *S. pombe* only Ser5 is absolutely essential and it can be compensated by tethering the capping enzyme to the CTD (17). Thr4 or Ser7 substitution with Ala are viable in both yeast species, but human and chicken cells harbouring Thr4 to Val mutation, or Ser7 to Ala in humans are lethal (17, 43, 44, 60, 61). In contrast, substitution of Ser7 by the phospho-mimic amino acid glutamate is inviable in both human and yeast cells (14, 61). Also insertion of additional amino acids between single heptads is lethal in *S. cerevisiae*, while insertion after a di-heptad is tolerated suggesting that it is the minimal functional unit of the CTD (62, 63).

1.1.2 Other mechanisms coupling transcription-related processes

Pol II transcription is a highly regulated process, comprising sequentially occurring steps that include transcription preinitiation, initiation, promoter clearance, promoter-proximal pausing (in higher eukaryotes), elongation, termination, and reinitiation. These events and RNA processing are interconnected in a bidirectional manner (6). For example, initiation is necessary for efficient pre-mRNA capping, which in turn influences downstream transcription and processing steps via the action of Capping enzymes (CEs) as well as CBC (see 1.2.1 and 1.2.2). Such proteins with dual function, involved in transcription and co-transcriptional processing, are the key players in coordination of these events. On the other hand, RNA processing steps depend also on the promoter structure, which is mediated through the recruitment of different factors and through alterations of the Pol II transcription rate (6).

The rate of transcription impacts co-transcriptional processes, which in turn can affect transcription. Pre-mRNA splicing, especially in higher eukaryotes, is regulated by the speed of the elongation complex (6). Especially, the suboptimal splicing signals require Pol II to slow down within the exon to be efficiently recognized by the spliceosome. Multiple factors modulate Pol II kinetics including regulatory DNA elements or phosphorylation status of the CTD. Pol II can be slowed down or even paused to give time for processing apparatus to perform their tasks (64). Pol II pausing was shown to play a key role in capping and transcription termination, where it acts as an especially important factor for the choice between alternative PASes (see 1.2.1 and 1.3.1). Also NNS-dependent termination is sensitive to Pol II kinetics (see 1.3.2).

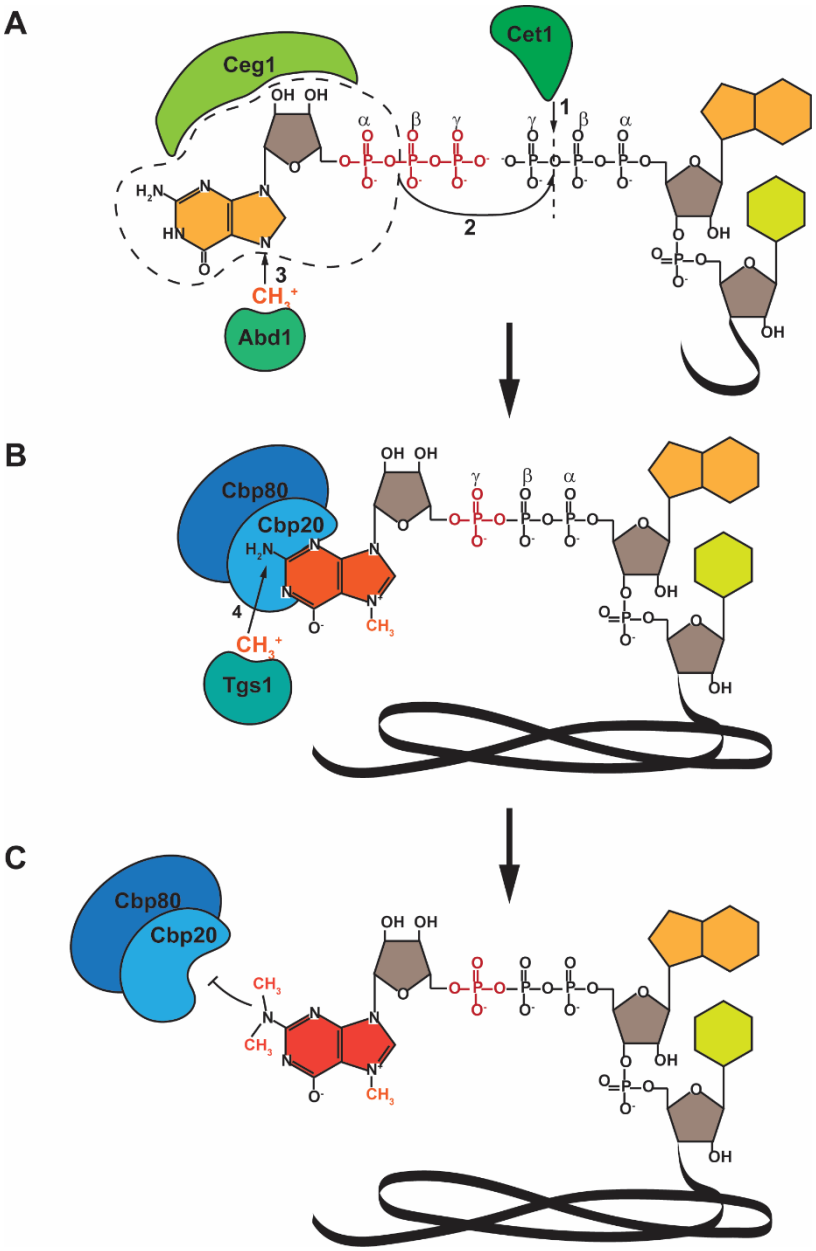

Figure 2. m^{7}G and TMG cap: structure and synthesis. *(**A**) Capping reaction: RNA dephosphorylation catalysed by Cet1 (1). Ligation with the inverted GMP molecule catalysed by Ceg1 (2). Added guanosine methylation at N7 position catalysed by Abd1 (3). (**B**) Cap recognition by CBC, composed of Cbp20 and Cbp80 subunits; m^{7}G dimethylation at the N2 position catalysed by Tgs1 (4). (**C**) Low affinity of CBC to the TMG cap. Lines with arrowheads indicate catalysed reaction, lines with barheads - inhibition. RNA is drawn as a black ribbon. Description in the text.*

Another factor that imposes a strong impediment to transcription elongation kinetics is chromatin structure, regulated by numerous highly conserved proteins. These proteins include chromatin remodeling enzymes, histone modification enzymes and histone chaperones and are also often recruited by the properly modified Pol II CTD (65). Elevated histone acetylation level can alter the rate of Pol II elongation in yeast cells with mutations in the Paf1 complex (polymerase-associated factor 1) that are deficient in H3 K4 and K36 methylation (66, 67). Dysfunction of Paf1 was also shown to lead to transcriptional readthrough at several

snoRNA genes, as Paf1 seems to facilitate Nrd1 recruitment, in addition to its role in histone methylation (67–69). Other transcriptional regulators, such as Rad6 required for monoubiquitination of histone H2B K123, and Set2 methyltransferase that trimethylates H3 K36, can affect NNS-dependent transcription termination in a locus-specific manner (69). Also Set1 is an important regulator of the NNS recruitment, it affect Pol II kinetics in early elongation by modulating histone acetylation (70). In turn, chromatin remodelers such as Isw1 and Chd1, were shown to prevent readthrough of PAS on mRNA genes (71, 72).

1.2 Capping and 5’ end maturation

1.2.1 m^{7}G cap formation

m^{7}G cap, that is a unique feature of all Pol II primary transcripts, is synthesized co-transcriptionally at the early stage of transcription (73, 74). This substantial modification has been shown to be required for viability in different organisms, reflecting its fundamental role in many cellular processes, including translation initiation. Three-step capping reaction is initiated when the 5’ end of the nascent transcript emerges from the Pol II RNA exit channel and is completed before the growing RNA chain reaches 100-150 bp (33). In the first step, terminal γ-phosphate group is removed from the 5’ triphosphate end by RNA triphosphatase Cet1 (Fig. 2A) (74). Resultant diphosphate is then ligated with the inverted GMP moiety by guanylyltransferase Ceg1 forming a unique 5’-5’ triphosphate bridge (Fig. 2A). Finally, the cap methyltransferase Abd1 specifically methylates the added guanine in the N7 position, using *S*-adenosylmethionine (AdoMet) as a donor (Fig. 2A). In yeast *S. cerevisiae*, all three reactions are catalysed by separate enzymes, whereas in metazoans first two reactions are performed by a bifunctional capping enzyme.

Capping is not only an important processing reaction, its completion serves as a signal for initiation-elongation transition of Pol II (6). It is therefore precisely synchronized with the transcription cycle. Capping enzymes are recruited to Ser5-P CTD and recent studies in yeast showed that this process occurs in a consecutive manner (33). Cet1-Ceg1 heterodimer is bound first and it dissociates shortly after the 5’-5’ triphosphate linkage synthesis. The subsequent recruitment of Abd1 is also mediated by Ser5-P CTD. In turn, Abd1, Ser5-P and cap methylation itself stimulate recruitment of Ser2 kinases Ctk1 and Bur1. Additionally, *S. cerevisiae* capping enzymes can also bind the Rpb1 ‘‘foot domain’ (75). The mammalian capping enzyme interacts not only with Ser5-P but also with Tyr1-P of the next CTD repeat (76). CTD Ser2-P promotes release of capping enzymes and facilitates events

leading to the assembly of the productive elongation complex. In addition, Cet1 was shown to promote this transition, independently of its mRNA capping activity or cap-mediated recruitment of CBC, as it decreases the accumulation of Pol II towards the promoter proximal site (77). On the other hand, Cet1 inhibits reinitiation which appears to provide sufficient time for capping and may serve as a quality control mechanism (78). Additional surveillance mechanism that ensures proper Pol II transcripts capping involves decapping by Rai1 and exonucleolytic degradation by Rat1 (see 1.2.4.2 and 1.2.5) (79, 80).

1.2.2 m^7G cap and CBC functions

Immediately following its synthesis, m^7G cap is recognized by the nuclear cap binding complex (CBC), a heterodimer composed of Cbp20 that directly binds the cap structure and Cbp80 that facilitates this interaction (Fig. 2B) (81–83). Neither Cbp20 nor Cbp80 can bind alone to the capped RNA, and moreover Cbp20 is unstable in the absence of Cbp80, but the opposite is not the case (84–87). CBC is recruited directly to the capped transcript through the highly specific interaction with the m^7G base, but it also interacts with the RNA chain through the RNA recognition motifs (RRM) in Cbp20 and Cbp80 (33, 86–88).

CBC binding stabilizes m^7G capped RNAs and is involved in regulation of co-transcriptional pre-mRNA maturation, its nucleocytoplasmic export and transcription itself (Fig. 3) (reviewed in (73, 74)). It enhances the rate of recognition and splicing of the cap-proximal intron in higher eukaryotes and is necessary for co-transcriptional spliceosome assembly at intron-containing genes in yeast (81, 89–92). Cbp20 was identified as a component of the yeast commitment complex stabilizing binding of U1 snRNP to the 5’ splice site and Mud2 to the intron branch point (Fig. 3) (84, 91, 93). In higher eukaryotes, CBC also regulates pre-mRNA 3’ end formation and its depletion from HeLa cell nuclear extract reduces the endonucleolytic cleavage step during this process (90, 94). 3’ end processing of capped noncoding and coding RNAs, including histone mRNAs, was demonstrated in humans to be stimulated also by a larger CBC-containing complex, CBCA, that additionally comprises ARS2 (arsenic-resistance protein 2) (95, 96). The interaction of ARS2 with transcription termination or 3’ end processing factors, such as CLP1, NELF-E, SLBP15 or FLASH, enables CBCA to facilitate recruitment of these factors (see 1.3) (95, 96). Also yeast CBC contributes somehow to mRNA 3’ end formation by impeding recruitment of cleavage factors complex CF IA to weak terminators and reducing termination and poly(A) addition at these sites (Fig. 3) (86).

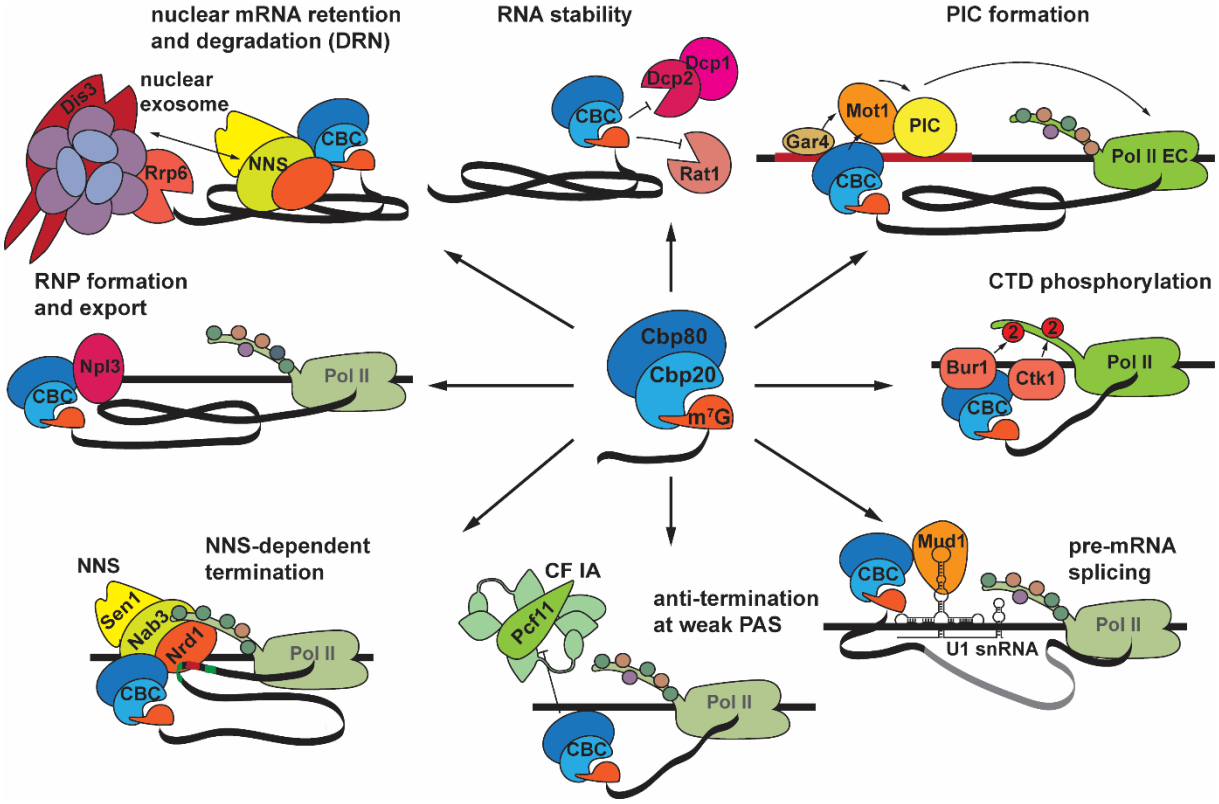

Figure 3. CBC functions in yeast. *CBC interacts with a spectrum of factors mediating RNA metabolism and stability, Pol II CTD phosphorylation, transcription termination and reinitiation. Lines with arrowheads indicate either activation, catalysed reaction or interaction (arrowheads at both ends), while lines with barheads indicate inhibition. Circles on the Pol II tail (CTD) indicate phosphorylated residues and are colour-coded as on Figure 1. RNA is drawn as a black ribbon, important sequences are coloured in grey (intron), green (Nab3 binding sites) and red (Nrd1 binding site). DNA template is drawn as a black line, red line indicates promoter region. Description in the text.*

CBC is also important for transcription itself. In yeast it interacts with Mot1 (modifier of transcription) that is recruited to gene promoters and enhances the formation of the preinitiation complex (PIC) in the subsequent rounds of transcription initiation (Fig. 3) (97). Recent studies showed that CBC facilitates Pol II CTD Ser2 phosphorylation (see 1.1.1) and thus contributes to proper Pol II elongation and alternative splicing, through its interaction with P-TEFb in mammalian cells or Bur1/2 and Ctk1/2/3 complexes in yeast (Fig. 3) (31–33).

Another cellular process that involves CBC is RNA degradation. In mammals, CBC is important for the recognition of the premature stop codon (PTC) in specialized mRNA quality control mechanism, Nonsense-mediated Decay (NMD), via direct interaction with a major NMD effector, Upf1, and stimulation of its binding to Upf2 (98). CBC is also a key component of the mRNA degradation pathway in the nucleus (DRN). The lack of Cbp80 stabilizes mRNAs, that are retained in the nucleus due to abnormal physiological conditions and are normally rapidly degraded through the pathway involving also Rrp6, exosome and Rat1 (Fig. 3) (99–101).

Recently a human CBCN complex, composed of CBC, ARS2 (that forms the CBCA subcomplex), ZC3H18 and the Nuclear Exosome Targeting (NEXT) complex, was described to link CBC to the 3'-5' RNA degradation, acting as a platform for exosome recruitment (95, 96). This complex was suggested to be important for the suppression of noncoding RNA transcription in human cells (see 1.3.2). The CBCN complex was also demonstrated to take part in biogenesis of human telomerase RNA (hTR) that contains a box H/ACA motif. This process is supposed to entail a kinetic competition between RNA processing and degradation by the exosome (102). Functionally similar CBC connection to the exosome may also exist in yeast *S. cerevisiae*, where CBC copurifies with the ncRNA-specific NNS transcription termination complex and the exosome (Fig. 3) (103). Similarly to human CBCN and its subcomplex CBCA, CBC-NNS connection can be involved in ncRNA transcription termination and 3' end formation.

CBC activity is regulated by interaction with importins that is probably mediated by growth factors and is important for temporal and spatial coordination of cap-dependent processes such as mRNA translation (74). In vertebrates importin-α binds Cbp80 through its N-terminal region and the CBC-importin-α complex stimulates the release of capped RNA via displacing Cbp20 from its contact with Cbp80 (104). This tuning of CBC-cap interaction is regulated by GTPase Ran that while bound to GTP disables the interaction of CBC-importin-α with importin-β. Thus, CBC-importin-α remains attached to the capped RNA in the nucleus, where the concentration of Ran-GTP is high, and is released in the cytoplasm where Ran-GTP level is low. CBC-importin-α-importin-β can be then imported back to the nucleus and reused. CBC interaction with importin-α is stable within the nucleus through multiple stages of RNA processing and is well conserved among eukaryotes (74).

Despite CBC important roles in diverse biological processes, yeast CBC is not essential for viability, and deletion strains show moderately reduced growth phenotype (84). Because only the heterodimer have a cap-binding ability, *cbp20Δ* and *cbp80Δ* show similar phenotypes, but surprisingly, the double *cbp20Δ cbp80Δ* mutant grows better than single deletion strains, which suggests that expression of either *CBP80* or *CBP20* alone have a dominant negative effect. The lack of CBC causes synthetic lethality with deletion of U1 snRNP components or other splicing factors, in agreement with the role of CBC in splicing (84, 105). Synthetic lethal interactions with CBC were also described for *CBF5* and *NOP58*, encoding components of H/ACA and box C/D snoRNPs, respectively, although the meaning of this genetic interaction is not known. CBC mutants show rRNA processing defects, possibly as a

result of splicing defect of a subset of ribosomal protein pre-mRNAs (84). Moreover, hypomorphic mutations of Cbp20 cap binding pocket (as *cbc2-NΔ42*) are synthetically lethal or sick with null mutations of factors involved in early spliceosome assembly and severely impede yeast sporulation and meiosis, as CBC is especially important for splicing of meiotically-expressed *MER3* and *SAE3* mRNAs (106). Another hypomorphic mutation, *cbc2-Y24A*, that attenuates CBC affinity for m^7G cap, completely suppresses the cold sensitivity of the *tgs1*Δ mutant (see 1.2.3) (107).

1.2.3 TMG cap

The m^7G cap of snRNAs, some snoRNAs and telomerase RNA component TLC1 is further dimethylated by Tgs1 to the 2,2,7-trimethylguanosine cap (TMG) (Fig. 2B) (108–110). Tgs1 is a conserved protein that catalyses two successive methyl transfers from AdoMet to the N2 position of m^7G. Its activity is strictly dependent on prior N7 methylation, which provides modification of properly capped RNA (111). Although Tgs1 catalytic activity does not require additional cofactors (111), Tgs1 is recruited to its substrates through interactions with sn/snoRNP specific proteins SmB and SmD1 (snRNA), Cbf5 (box H/ACA snoRNA) and Nop58 (box C/D snoRNA) (108). Another binding partner of Tgs1 that is involved in snRNA and U3 snoRNA cap trimethylation, acting as a Tgs1 specificity factor is Swm2 (112).

In mammals, snRNP cap hypermethylation occurs in the cytoplasm (74, 109). The m^7G cap bound to CBC acts as an export signal, recognized by the snRNA-specific export adaptor protein PHAX and transported to the cytoplasm via the export receptor CRM1. Following the Sm-core assembly, cap hypermethylation and 3' end trimming in the cytoplasm, snRNPs are imported back into the nucleus using the TMG cap and Sm-core as the nuclear localization signal and Snurportin-1 and the survival motor neuron (SMN) complex as adaptors (see 1.5) (109, 113). Cap hypermethylation of those mammalian snoRNAs that retain this structure is believed to occur in Cajal bodies (see 1.5), whereas in yeast m^7G cap of both snoRNAs and snRNAs is probably modified in the nucleolus, although it was not conclusively shown, despite Tgs1 localization in this structure (108, 114). TMG-capped RNAs do not further bind CBC whose affinity to 2,2,7- trimethylguanosine is much lower than to m^7G (Fig. 2C) (115).

Mammalian TMG cap of mature snRNPs is essential for their nuclear localisation, but yeast sn/snoRNPs do not transit through the cytoplasm during their biogenesis (109) and neither their TMG cap nor Tgs1 are required for cell viability (107, 108). The composition of sn/snoRNPs in the absence of TMG cap is only slightly altered, with CBC still attached to their m^7G (107). *S. cerevisiae tgs1Δ* mutant shows a cold-sensitive splicing phenotype, related

to the retention of U1 snRNA in the nucleolus, which is suppressed by the mutation in the cap-binding pocket of Cbp20 (*cbc2-Y24A*) (107). *TGS1* deletion causes synthetic lethality with genes encoding pre-mRNA splicing factors such as Mud2 and Nam8, and a weaker synthetic growth defect with a number of genes coding proteins involved in snRNP function and spliceosome assembly, including Brr1, Lea1, Ist3, Isy1, Cwc21, and Bud13 (116). This suggests that TMG cap is involved in spliceosome assembly and is functionally redundant with proteins that assist U1 and/or U2 snRNP recruitment to the pre-mRNA. Although Tgs1 is not essential for splicing in mitotically growing yeast, it is required for yeast sporulation, promoting splicing of meiotic pre-mRNAs *PCH2* and *SAE3* with non-canonical splicing signals (117).

It was also suggested that TMG cap may play some role in snRNA 3' end processing, because of a strong synthetic sick interaction between *TGS1* and *TRF4*, encoding a catalytic subunit of the TRAMP complex that cooperates with the exosome in sn/snoRNA 3' end formation (see 1.3.3.3) (116). A synthetic sick phenotype was described for *TGS1/LSM1* and *TGS1/PAT1* pairs. Lsm1 and Pat1 are both involved in RNA decapping that may contribute to 5' end processing of some snoRNAs (see 1.2.5).

Tgs1 is also required for pre-rRNA processing and for maintaining nucleolar morphology (118) and in addition its absence is associated with changes in telomere length and structure, resulting in shorter replication lifespan and premature aging (110).

1.2.4 Endo- and exonucleolytic 5' end processing

Part of snoRNA precursors, more commonly of the box C/D class, undergo further 5' end processing that removes the cap structure and thus also CBC from pre-snoRNPs. 5' end of these pre-snoRNAs is extended and contains a steam-loop structure that is recognized and cleaved by Rnt1 endonuclease (119, 120). This provides an entry site for further exonucleolytic trimming catalysed by Rat1 and Xrn1 nucleases (see 1.5).

1.2.4.1 Rnt1

Rnt1 belongs to the RNase III endonuclease family that specifically cleaves a double-stranded RNA (dsRNA) introducing two staggered cuts on each side of the RNA duplex (121). It is composed of dsRNA binding (dsRBD) and nuclease domains conserved among bacterial and eukaryotic RNase III family, and the N-terminal domain that promotes enzyme homodimerization and increases the accuracy and efficiency of its catalytic activity (122, 123). Rnt1 substrate comprise a hairpin structure capped with a tetraloop containing the A/uGNN consensus sequence that is important for the recognition by the enzyme and defines

cleavage sites 14–16 base pairs from the tetraloop (Fig. 4) (123–125). Although the second nucleotide of the tetraloop is universally conserved and its mutation significantly reduces binding and blocks cleavage, AAGU tetraloop of snR48 is also recognized and cleaved by Rnt1 with a comparable affinity due to cooperative conformational changes in both the RNA and the enzyme (126–128).

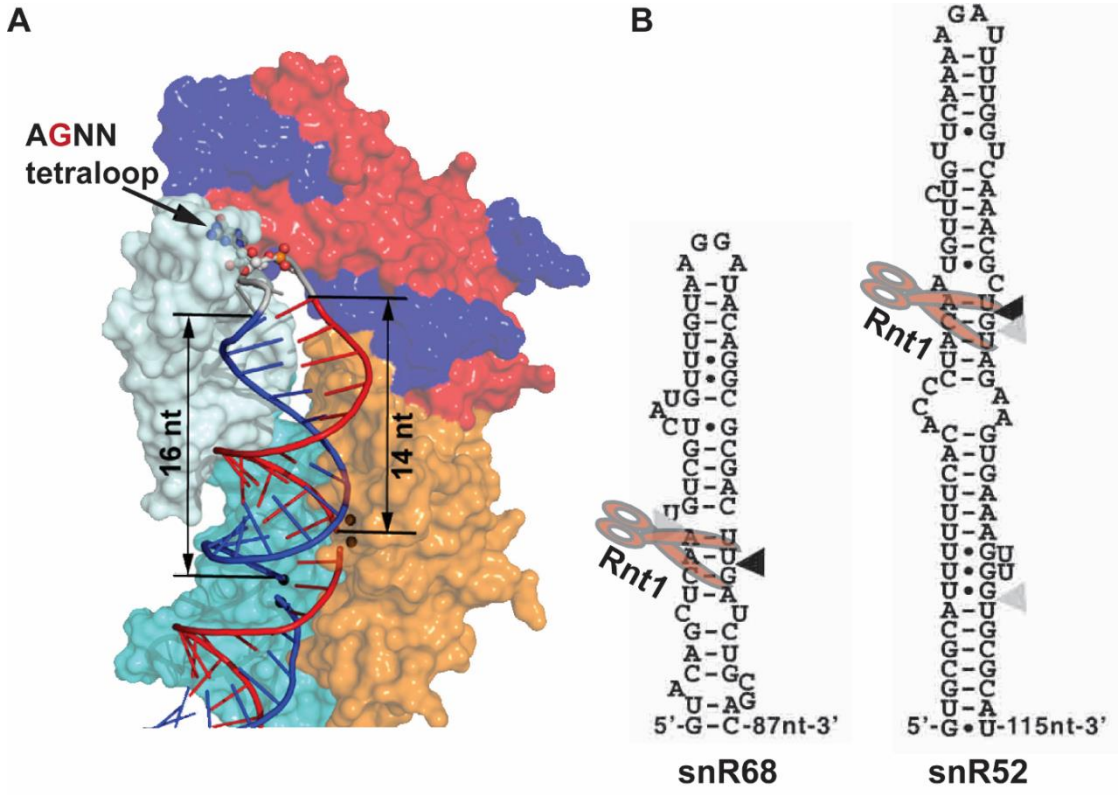

Figure 4. Rnt1 substrate recognition and cleavage site determination. *(**A**) Rnt1 dimer interacting with the stem terminated with the AGUC tetraloop, adapted from* (123)*. Protein is illustrated as a molecular surface, RNA is shown in grey for the tetraloop (guanine residue is highlighted as a ball-and-stick illustration) and in blue and red for the stem. (**B**) Secondary structure prediction of RNA sequences found upstream of box C/D snoRNAs, adapted from* (120)*. The distance to the mature snoRNA sequence is indicated on the 3' end of the structure. Cleavage sites mapped in vitro are indicated by black triangles and cleavage sites mapped in vivo are indicated by grey triangles.*

The A/uGNN tetraloops are commonly found in yeast snRNA, snoRNA and rRNA precursors processed by Rnt1 (119, 120, 129–132). In the case of pre-rRNA, Rnt1 co-transcriptional cleavage downstream of the mature 25S is strongly linked to transcription, as Rnt1 is recruited to the nucleolus in the presence of actively transcribed rRNA genes and physically interacts with RNA polymerase I (Pol I) (132, 133). In contrast, the co-transcriptional or post-transcriptional nature of Rnt1-mediated processing at the 3' end of yeast snRNAs or 5' end of snoRNAs has not been explicitly demonstrated. However, the role of Rnt1 in pre-sn/snoRNA processing was reported to be facilitated by its physical

interactions with snoRNP proteins Nop1 and Gar1 or termination factors Sen1 and Nrd1 (see 1.3 and 1.5) (36, 103, 134, 135).

Rnt1 also contributes to the degradation of some unspliced pre-mRNA precursors by cleaving intronic stem-loops present in unspliced precursors and lariat introns (136). In the case of rare yeast genes with two introns, such as a modulator of sex-specific gene expression *MATa1*, Rnt1 eliminates pre-mRNAs folded in a manner that promotes the skipping of the second exon (137). Rnt1-dependent nuclear degradation is involved in regulation of expression of specific mRNAs under distinct environmental conditions, including telomerase subunit *EST1* mRNA, mRNAs encoding iron uptake and iron mobilization proteins or cell wall integrity associated mRNAs (138–140) In addition, Rnt1 cleavage of nascent transcripts mediates fail-safe transcription termination (see 1.3.3) (141, 142).

Independently of its endonuclease activity, Rnt1 is required for cell cycle progression and nuclear division. Rnt1 resides in the nucleolus in G1 and S phases and its cell cycle-dependent relocation to the nucleoplasm contributes to the exit from the G2/M phase (133). This cell cycle regulated control of Rnt1 localization could be responsible for the degradation of the telomerase *EST1* or cell wall integrity associated *HSL1* mRNAs in the G2/M phase.

Surprisingly, *RNT1* gene is not essential, although its deletion causes severe growth defect due to a globally disturbed metabolism of ncRNAs (143). Biogenesis of rRNA and many sn/snoRNAs is strongly disturbed in a temperature sensitive *rnt1*Δ strain, however, the level of some snoRNAs processed at their 5’ end (e.g., snR52) appears to be hardly affected (120). *RNT1* deletion also impacts the accuracy of Pol I termination (144, 145) and reduces rRNA transcription, at the same time increasing the number of transcriptionally poised rRNA genes (146). The growth defect in the absence of Rnt1 is also related to impaired cell cycle progression due to abnormal nuclei positioning (133). Moreover, upregulation of mRNAs associated with morphogenesis and cell wall integrity in cells lacking Rnt1 leads to stress response attenuation (140). Consistently, *RNT1* deletion shows strong synthetic sickness or lethality with cell wall integrity pathway related genes, and these phenotypes can be partially rescued by increased osmotic strength of the growth medium.

1.2.4.2 Rat1 and Xrn1

Cap structure contains a triphosphate bridge that is atypical for nucleic acids and cannot be cleaved by exonucleases. Therefore, maturation of extended precursors requires additional endonucleolytic cleavage or decapping steps that provide monophosphate 5’ end as a substrate for 5’-3’ exonucleases. Two major yeast 5’-3’ exonucleases, nuclear Rat1 and

predominantly cytoplasmic Xrn1, participate in 5’ end processing of different RNA species, with a more prevalent role of the nuclear Rat1 (147–151). The two enzymes are functionally equivalent and are fully interchangeable when targeted to the non-native cellular compartment (152–154). However, *RAT1* is essential for growth and *rat1-1* phenotypes are not rescued by *XRN1* overexpression, whereas *XRN1* is nonessential and elevated expression of *RAT1* partially restores growth and RNA turnover defects of *xrn1* mutants (152, 155–157)). While the main role of Xrn1 is in cytoplasmic mRNA turnover, Rat1 functions in the nuclear 5’ end-processing of ncRNAs and quality control of aberrant or inefficiently processed RNAs (158–160).

Rat1 in yeast interacts with a Rai1 cofactor that stabilizes and enhances its activity, assists in degrading structured RNAs and provides the access to monophosphate 5’ end when necessary (80, 154, 161). Besides the 5’ end formation of many snoRNAs, Rat1 is required for the maturation of the 5’ end of 25 S and 5.8 S rRNAs and it is involved in Pol I rRNA transcription and co-transcriptional cleavage by Rnt1 (158). Another important function of Rat1 is in transcription termination of Pol I and Pol II, at least partly by contributing to co-transcriptional degradation of nascent RNA in a so-called torpedo mechanism (see 1.3.1) (162, 163). Moreover, Rat1 enhances recruitment of 3’ end processing factors (164) and influences the phosphorylation levels of Pol II CTD (38). Rat1/Rai1 associate with the Pol II EC, partially through the interaction with the Ser2-P-CTD binding protein Rtt103, from the very beginning of the active transcription unit, peaking at the 3’ end (162). Rat1/Rai1 is also involved in premature termination of Pol II engaged in the production of uncapped or improperly capped RNAs as a part of cap-related quality control mechanisms that prevent incomplete mRNAs from being transported and translated (see 1.2.5) (79, 80). In fact, Rat1/Rai1 is the major component of prominent nuclear ncRNA surveillance pathways, responsible for the identification and degradation of aberrantly processed pre-rRNAs, antisense RNAs (asRNAs) and structurally compromised mature tRNAs as well as normal mRNAs retained in the nucleus (159). In addition, Rat1 associates with telomerase and degrades telomeric repeat-containing ncRNA TERRA, thus promoting telomere elongation via a mechanism mediated by the repressor-activator protein Rap1 and its interacting proteins, Rif1 and Rif2 (165, 166). Rat1 is also involved in promoting proper replication fork progression (165).

The *rat1-1* temperature-sensitive allele was identified in a screen for factors involved in the export of poly(A) mRNA from the nucleus, although this phenotype was not observed for

another *RAT1* allele (152). It was also shown to cause the accumulation of the ncRNA TERRA that inhibits telomerase activity due to RNA/DNA hybrid formation (165). This defect results in telomeres shortening and can be overcome by RNaseH overexpression. The *rat1-1* mutation causes elevated CTD phosphorylation, altered Pol II distribution and, most importantly, transcription termination defects (38). These phenotypes are rescued by overexpression of the CTD phosphatase Fcp1. Moreover, the *rat1-1* allele is responsible for increased Pol II transcription kinetics and therefore is synthetic lethal with a mutation in Rpb1 that increases Pol II speed and suppressed by mutations slowing polymerase down (38, 79). On the other hand, *rat1-1* rescues Pol II processivity defect. Interestingly, *rat1-1* is also synthetic sick with a mutation in guanylyltransferase Ceg1 due to the toxicity of uncapped mRNAs. (167).

1.2.5 Decapping

Decapping by cap-specific pyrophosphatases (decapping enzymes) is an important step of all RNA degradation pathways for Pol II transcripts (160, 168). Especially, as mentioned above 5'-3' exonucleolytic decay requires prior hydrolysis of the pyrophosphate bond.

The major cellular decapping enzyme engaged in these processes is Dcp2 from the Nudix pyrophosphatase family that acts within a dimeric complex with Dcp1 that is required for cap hydrolysis *in vivo* (169, 170). The Dcp1/Dcp2 complex preferentially acts on longer RNA substrates, at least 12 nucleotides long (171–174). Dcp2 Nudix pyrophosphatase domain is flanked by conserved boxes A and B that interact with Dcp1 and RNA substrate, respectively (172, 175, 176). Dcp1 stimulates Dcp2 activity, affecting reaction chemistry but not substrate binding, and serves as a bridge between the enzyme and decapping activators and inhibitors that regulate its activity *in vivo* (173, 175–177). In higher eukaryotes additional factors are required to stabilize the interaction between Dcp1 and Dcp2 such as Edc4/Hedls/Ge-1 (VARICOSE in Arabidopsis) (178, 179). Moreover, multiple enhancers, including yeast Dhh1, Pat1, the Lsm1-7 complex or Edc1-3, and inhibitors of decapping, such as eIF4E or poly(A)-binding protein Pab1, are involved in this process (169, 170). Some of them act as general factors inhibiting or facilitating the assembly of the decapping machinery or influencing Dcp2 enzymatic activity, and some are RNA-specific. Although active decapping machinery is localized mainly in cytoplasmic foci termed P-bodies it was shown that cap hydrolysis can occur co-translationally and the exact function of these foci remains unclear (171, 180–183). Decapping enzymes compete with cap binding proteins for the access to the cap structure, and binding of different cap enhancers leads to formation of the decapping

machinery and its association in P-bodies probably by interfering with translation initiation factors (160).

Although Dcp1/Dcp2 is primarily cytoplasmic it was shown to shuttle into nucleus where it can be involved in numerous RNA surveillance decay pathways that engage 5'-3' RNA degradation (see 1.2.4.2) (184, 185). Nuclear decapping is stimulated by the Lsm2-8 complex and it plays an important role in utilization of nucleus-restricted mRNAs (184) and in regulation of expression of lncRNAs that are degraded either by the Xrn1-dependent pathway in the cytoplasm or by Rat1 in the nucleus (186). The latter has further impact on transcriptional activity of inducible genes such as *GAL* regulon (186, 187). Yeast and human Dcp1/Dcp2 complex cleaves not only m^7G but also TMG and unmethylated capped RNAs and it was recently shown to be involved in the quality control of U1 snRNA in human and yeast cells (188–190). There is some evidence that Dcp1/Dcp2 as well as Xrn1 are present in the nucleus and associate with the chromatin to promote transcription (191). It is therefore possible that RNA decapping may occur co-transcriptionally in degradation of aberrant transcriptional products or 5' end exonucleolytic processing of snoRNAs. Moreover, it was recently shown to function in premature termination of human Pol II together with the Rat1 homolog, Xrn2, suppressing promoter bidirectionality (192).

Besides Dcp1/Dcp2 there are other less studied decapping enzymes responsible for hydrolysis of a subset of RNAs. Nudt16, initially identified in *Xenopus laevis* as a nucleolar enzyme involved in decapping of U8 snoRNA (193, 194), in mammalian cells is localized in the cytoplasm, where it contributes to mRNA decapping (195). In turn, the cofactor of Rat1 exonuclease, nuclear Rai1 has a decapping activity preferentially towards the unmethylated cap (GpppG) and also a pyrophosphatase activity, acting as an important player in RNA cap quality control (80, 161). Another decapping enzyme engaged in this process is Rai1 homolog Dxo1, that primarily resides in the cytoplasm and can degrade RNAs with an unmethylated cap single-handedly as it has both decapping and 5'-3' exonuclease activity (196). The mammalian homolog of these enzymes, DXO, performs all three activities and eliminates both uncapped substrates and RNAs with the unmethylated cap (197). In contrast to Rai1, which shows only minimal decapping activity on properly capped RNAs, Dxo1 and DXO efficiently hydrolyse m^7G structure (196, 197). Moreover, DXO also has a decapping activity towards TMG-capped RNAs, and may therefore be involved in degradation or processing of mammalian sn/snoRNAs (197). Recent studies identified six additional mammalian Nudix pyrophosphatases, namely Nudt2, Nudt3, Nudt12, Nudt15, Nudt17, and Nudt19 and the yeast

Nudt3, as decapping enzymes hydrolysing m^7G and unmethylated caps (189). Another cap specific pyrophosphatase conserved among eukaryotes is the decapping scavenger enzyme DcpS that acts on a free cap and short-capped RNAs that arise from 3'-5' mRNA degradation by the exosome (198). The important role of this enzyme is the utilization of free cap structures that sequester cap binding complexes from their RNA substrates (199).

Both Dcp1 and Dcp2 are not required for yeast viability, probably because their function overlaps with other decapping enzymes and the 3'-5' RNA degradation machinery. However, both *dcp1*Δ and *dcp2*Δ strains grow extremely slowly at all temperatures due to RNA decay defects (200, 201). Consistently, *dpc* mutants are synthetically lethal with mutations in genes involved in 3'-5' RNA degradation (202, 203).

1.3 Transcription termination and 3' end formation

The last step of the transcription cycle, termination, is tightly coupled to RNA 3' end processing and both processes are essential for the release of Pol II and the transcript from the template and each other (reviewed in (204–207)). However, transcription termination is also vital for general expression of the genome, as it hinders Pol II from interfering with downstream transcriptional units and prevents from formation of asRNAs that could interfere with the synthesis of normal cellular RNAs. This is particularly important in the light of recent evidence that eukaryotic genomes are promiscuously transcribed and this transcriptional traffic, as well as the fate of resulting products, must be tightly regulated (207). It was shown that defective termination may affect splicing efficiency and at the same time increase RNA degradation (208). Recent findings demonstrate that transcription termination and 3' end processing also impact initiation of the next round of transcription (209). This is connected with so called gene looping that brings promoter and terminator regions together, which in yeast is mediated by the interaction between TFIIB and CTD Ser5-P phosphatase Ssu72. Finally, as several yeast genes have been shown to undergo premature termination, this process may contribute to overall regulation of gene expression. Pol II termination is a dynamic process and can occur within, upstream or downstream of the transcription unit. Therefore, the timing of Pol II release influences the length of the final transcript, affecting its stability, localization and functionality.

Pol II termination can proceed through different mechanisms, depending on RNA 3' end processing signals and termination factors present at the end of the gene (205). The best-studied pathways are the poly(A)dependent termination, utilized in the case of protein coding genes, and mechanism employed mainly in ncRNA transcription in yeast that depends on the

complex of two RNA binding proteins Nrd1 and Nab3 and helicase Sen1 (NNS complex). Both mechanisms involve shared factors and are probably utilized as the main or fail-safe pathways, depending on the terminator and the context of other *cis*- and *trans*-acting factors (206, 207, 210). All events described below are highly regulated and coordinated through transcription cycle by binding and release of multiple factors that specifically recognize different elements in the Pol II complex. The coordination of these processes, especially by Pol II CTD modifications, is described in another chapter (see 1.1.1).

1.3.1 Protein coding genes: Cleavage and polyadenylation

Pol II termination downstream of most protein coding genes is mediated by the cleavage and polyadenylation (CP) machinery and the poly(A) signal (PAS) that acts as a 3' end processing and transcription termination site (Fig. 5A) (205–207). Human PAS consists of a highly conserved AAUAAA sequence and other less conserved elements, while *S. cerevisiae* PAS consensus is less conserved and in the case of most gens is composed of an adenine-rich efficiency element (EE, TAYRTA, where Y is any pyrimidine and R any purine), an A-rich positioning element (PE, AAWAAA, with W being adenine (A) or thymine (T)), located 30 nt upstream of the cleavage position, and an uridine-rich element spanning the cleavage and poly(A) site (206).

Transcription through the PAS followed by Pol II pausing allow for the endoribonucleolytic cleavage of the nascent transcript, subsequent polyadenylation of the upstream cleavage product and degradation of the downstream RNA (205–207). The strength of the PAS and the presence of many *trans*-acting termination factors influence the efficiency of termination. The machinery that carries out this reaction comprises more than 20 proteins in yeast and over 80 in mammals, which are grouped in three complexes (Fig. 5B): cleavage and polyadenylation factor (CPF) and cleavage factors IA (CF IA) and IB (CF IB) (205, 206). They are recruited to the Pol II complex via multiple interactions involving Pol II body, CTD and the nascent transcript, bound by termination factors at the moment PAS is transcribed by elongating Pol II (206). In yeast, binding of Rna15 (CF IA) to the PE and Hrp1 (CF IB) to the EE defines the cleavage site (Fig. 5B). Other components of CF IA, Rna14, Pcf11 and Clp1, are essential for termination and 3' end processing and provide the scaffold for the CPF recruitment. CPF is a large complex consisting of few subcomplexes: CFII (Pta1, Yhh1/Cft1, Ydh1/Cft2, Ysh1/Brr5), PFI (Mpe1, Pfs2, Fip1, and Yth1) and Pap1, forming the so called "core CPF" that dynamically associates with the APT complex (Pti1, Swd2, phosphatases Glc7 and Ssu72, Ref2, and Syc1) through the interaction with Pta1 (211). The latter is

considered to be a fine-tuner of the 3' end processing reaction (206). CPF associates with Pol II CTD via Ydh1 and Yhh1 that also interacts with RNA, and CF subunits, and binds the nascent transcript downstream of the cleavage site. Two CPF subunits, Ydh1 and Ysh1, contain βCASP family endonuclease domain, although only that of Ysh1 has been shown to be functional (212). The major yeast poly(A) polymerase Pap1 is recruited to the PAS through the interactions of Yth1 with Fip1 or directly with Pta1 (Fig. 5B and C) (206). Polyadenylation of pre-mRNA is regulated by poly(A) binding proteins Nab2 and Pab1 in yeast or PABII in mammals that adjust the poly(A) tail length to around 70 or 200 nt, respectively.

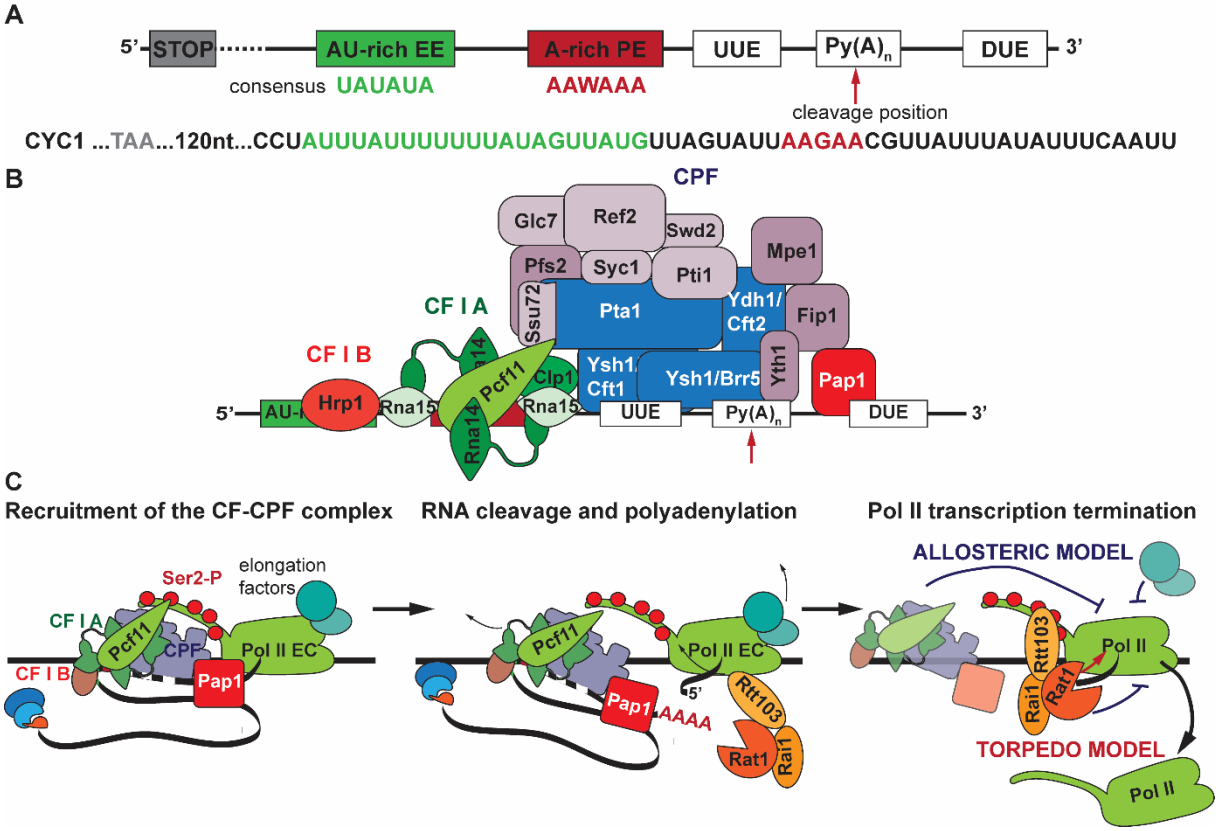

Figure 5. Polymerase II transcription termination at protein coding genes in yeast S. cerevisiae. *(**A**) CPF/CF binding sites: the AU-rich efficiency element (EE), the A-rich positioning element (PE) and U-rich regions (UE) surrounding the cleavage site (upstream and downstream UE), the cleavage site is indicated by a red arrow and is located within the A-rich region (Py(A)$_n$). Polyadenylation signal of CYC1 mRNA is shown to exemplify the low conservation of yeast CPF/CF-dependent terminators. According to* (206) *(**B**) Suggested stoichiometry and organisation of the CPF/CF complex on mRNAs, adapted from* (206). *Sub-complexes are colour-coded: blue-purple – CPF with APT (light purple), PFI (purple) and CFII (blue) subcomplexes, green – CF I A and red – CF I B (Hrp1). (**C**) Unified allosteric-torpedo model of Pol II transcription termination at protein coding genes, based on* (207). *Lines with arrowheads indicate either activation or complex dissociation, while lines with barheads indicate inhibition. Circles on Pol II CTD tail indicate phosphorylated residues and are colour-coded as on Figure 1. RNA is drawn as a black ribbon, terminator elements and termination factors are coloured as on the schemes presented above. DNA template is drawn as a black line. Description in the text.*

Although Pol II transcription termination is tightly coupled to 3' end processing and mutation of many processing factors as well as mutation in the PAS cause termination

defects, Pol II continues transcription after cleavage to finally dissociate from DNA even up to several hundred bases downstream of the PAS (206). Transcription of the PAS reduces Pol II association with the template probably by decreasing the speed of the EC (Pol II pausing) and destabilising the RNA:DNA hybrid within polymerase catalytic centre (206). The cleavage reaction *per se* is important for Pol II release, but the exact mechanism of polymerase dissociation from the DNA template is not fully understood and two models of termination were proposed (207). According to the 'torpedo' model, a free monophosphate 5′ end generated by the 3' end processing reaction is recognized and rapidly degraded by Rat1 that quickly catches up with Pol II and promotes its release from the template, possibly via direct collision (213). On the other hand, a second mechanism, known as the allosteric or anti-terminator model, attributes polymerase release to allosteric changes in the EC that decrease Pol II processivity and destabilize its interaction with the DNA template (214). These changes are caused by dissociation of some elongation factors and/or association of termination factors. There is some evidence that these mechanism are not mutually exclusive, as the torpedo action of Rat1/Rai1 is not sufficient to terminate Pol II and mutations of several CP factors, including Rna15, Rna14, Pcf11, Yhh1, Ysh1, and Ssu72, result in readthrough at the 3' end of protein coding genes (164, 212, 215–217) . A complete termination mechanism more likely reflects a combination of both strategies in a unified allosteric-torpedo model (Fig. 5C) (164, 218–220).

There are conflicting data concerning the role of the exonuclease activity in transcription termination. It seems not to be sufficient for Pol II release as the nascent transcript degradation by Xrn1 redirected to the nucleus does not promote termination (164). In addition, catalytically inactive Rat1 (*rat1D235A*) is able to dismantle the EC *in vitro* upon addition of Rtt103 (221). On the other hand, recent findings demonstrate that Rat1/Rai1 is sufficient to terminate Pol II *in vitro* in the presence of nucleoside triphosphates (222). Rat1 exonuclease activity is essential for this process, as the *rat1* catalytic mutant fails to terminate Pol II, regardless of the presence of Rtt103 (222, 223). The termination action of Rat1/Rai1 is attributed to Pol II pausing caused by NTP misincorporation that provides time to degrade nascent RNA and 'torpedo' the polymerase. Recent studies in human cells also suggest that the kinetic competition between Pol II and Rat1 homologue Xrn2 is important for proper transcription termination on most human genes (224). It transpires that Xrn2 dominant negative catalytic mutation delays Pol II termination genome-wide, whereas Pol II with altered elongation rate (slow/fast) accordingly impacts the occurrence of termination

(early/late). On the other hand, mammalian native elongating transcript sequencing (mNET-seq) showed, that Xrn2 is involved in termination of sense and antisense promoter-associated non-productive transcripts and its depletion does not cause general termination defects (11). In line with the latter, PAS-dependent termination of mammalian Pol II was shown *in vitro* to involve only PAS recognition leading to Pol II pausing followed by at least two conformational changes, but not to require RNA cleavage or 5'-3' exonuclease activity (225). However, cleavage-dependent and independent torpedo mechanisms are not mutually exclusive and can operate in parallel.

Specific interaction between Rat1 and Pol II, possibly aided by Rtt103, was proposed to assist in the transmission of the termination signal (221, 222). Nevertheless, the role of Rat1 in Pol II termination seems to be far more complicated than proposed in the 'torpedo' model as it also impacts the CTD phosphorylation status and Pol II kinetics (see 1.2.4.2) (38, 79).

Binding of many protein factors to the elongating Pol II makes it competent for termination (204–207). The most important for efficient termination is Pcf11, a CF IA component that associates with Ser2-P CTD and is involved in the cleavage reaction (226–228). Independently of this function, it seems to bridge the CTD to pre-mRNA and dismantle the elongation complex thereby coupling 3' end processing to termination (227, 229). Interestingly, *in vitro* studies and uncoupling of these two steps in *pcf11* mutants *in vivo* showed that CTD binding is in fact important for proper termination but not for 3' end formation (215, 223, 227, 230). Moreover, depletion of human Pcf11 reduces the efficiency of Pol II termination stabilizing the 3' cleavage product, suggesting its role in Xrn2-dependent degradation of the nascent transcript (231). In addition to Pcf11-CTD interaction, Pcf11/Clp1 subcomplex associates with the polymerase body through a flap loop domain of Rpb2 which is important for transcription termination of short genes (see 1.3.2) (232). In yeast, Pcf11 and Rat1 appear to act in a cooperative manner in facilitating reciprocal recruitment to the elongation complex (164).

Dissociation of other factors, so called anti-terminators, also influences EC conformation, thus favouring Pol II termination. Paf1C and the TREX/THO complex, implicated in coordination of mRNP biogenesis and maintenance of genome integrity, are associated with the transcription machinery throughout genes and their level is reduced past the poly(A) site (204). Yeast Sub1 and its human homologue PC4 have a conserved anti-terminator activity as their dissociation from the CP machinery upon PAS recognition facilitate termination (204).

Another anti-terminator that prevents premature Pol II termination is Npl3 that competes for RNA binding with CF I components Rna15 and Hrp1 (233, 234). Npl3 interacts with Ser2-P CTD, which increases Pol II elongation rate (235). This interaction is inhibited by Npl3 phosphorylation by Casein Kinase 2 (CK2), affecting Npl3 ability to compete against Rna15 for RNA binding. In turn, Npl3 methylation by an arginine methyltransferase Hmt1 stimulates its export from the nucleus and weakens interaction with Tho2, a subunit of the THO and TREX complexes (236, 237). Hmt1 was shown to promote elongation and to suppress termination at cryptic terminators and may facilitate recycling of Npl3 and Tho2 (238). The important function of Npl3 is to ensure the fidelity of pre-mRNA processing and mRNP formation prior to termination. Similarly, CBC that interacts with Npl3, was shown to independently act as an anti-terminator that prevents early termination at weak PASes (86). However, it does not influence Pol II elongation rate, but interferes with binding of Rna15 and Pcf11. Both proteins are also engaged in inhibiting premature termination at naturally occurring weak PASes within the *RNA14* gene, serving as an additional regulation system, as Rna14 stabilizes Rna15 and Hrp1 interaction with pre-mRNA and its overexpression increases recognition of weak PASes (234, 239). The mechanism of CBC suppression of cryptic PASes was not fully described and it is unclear whether it acts on its own or through interaction with other factors such as U1 snRNP. It was shown in mammals that U1 has a very strong anti-terminator activity and is implicated in regulating promoter directionality (240–243). Recognition of proximal PASes located in the sense orientation is inhibited by U1 snRNP binding, whereas U1 binding sites are depleted in the antisense orientation, where a high density of PASes promotes early termination, thus defining promoter directionality.

Other factors that play important roles in transcription termination of protein coding genes are Pol II subunits Rbp3 and Rpb11. They were proposed to “read” or “transfer” the termination signal, as their mutations cause readthrough at both mRNA and snoRNA terminators (244). Although these mutations are distant from Pol II RNA exit channel, it is possible that Rpb3/Rpb11 surface interacts with some termination factor such as APT subunits or Sen1 helicase. Pol II subunit Rpb4 is also involved in recruiting termination factors, including Rna14, to the EC (245). Moreover, loss of Rpb4 results in altered PAS usage at the *RNA14* gene as a part of a feedback upregulating Rna14 levels in response to a low CF IA recruitment. Rpb4, located near CTD of Rpb1 and the RNA exit channel, probably stabilizes interactions of 3’ end processing factors with CTD and PAS on the nascent transcript.

Sen1 (senataxin in human cells), the putative RNA/DNA helicase, was proposed to cooperate with Rat1, exposing the downstream cleavage product and facilitating its degradation (246). However, loss of Sen1/senataxin function does not affect termination at all poly(A)-dependent genes and its depletion does not result in readthrough at most protein coding genes (220, 247).

1.3.2 Noncoding genes: NNS-dependent transcription termination

Transcription termination of *S. cerevisiae* ncRNAs, such as sn/snoRNAs or cryptic unstable RNAs (CUTs), depends on a distinct mechanism employing the NNS complex as a key player (204, 206, 207, 248). The NNS complex comprises RNA-binding proteins Nrd1 and Nab3 and the putative RNA/DNA helicase Sen1. All three proteins and their roles are described in details in subsequent chapters (see 1.3.2.1-3). In addition, several mRNA 3' end processing factors have been shown to be involved in the transcription termination of Pol II transcribed ncRNAs (see 1.3.2.4). Their interaction with the elongation complex could facilitate some conformational changes, stimulating its dissociation from the DNA template. However, the major activity that dismantles Pol II complex is attributed to Sen1 (see 1.3.2.3) (Fig. 6C).

snoRNA terminators appear to be bipartite (Fig. 6A), with a major NNS-dependent terminator (terminator I) containing clusters of Nrd1 and Nab3 binding sites (see 1.3.2.1) and downstream mRNA-like signal (terminator II) (217, 244, 249–251). Both signals are required for efficient Pol II termination that generates two classes of polyadenylated pre-snoRNAs (251). The recognition of ncRNA terminator either by the NNS or CPF complexes was proposed to depend on the distance from the transcriptional start site (252), but recent studies show that both complexes probably cooperate in efficient termination of sn/snoRNA transcription (41). Notably, ncRNA termination in yeast growing at 25°C overlapped with Nrd1/Nab3 binding sites (terminator I), whereas at 37°C it was shifted downstream to distal regions (terminator II), where Pcf11 level reached its maximum. It was proposed that increased Pol II elongation kinetics at elevated temperature affects Sen1 termination activity (253). Higher concentration of Pcf11 on chromatin is therefore necessary to stimulate Sen1 (41). Another constituent of the non-poly(A) terminator is a region containing a series of U residue runs located downstream of I and II sequences (220). When Pol II fails to recognize these terminators it continues transcription until the next PAS (206).

sn/snoRNA transcription termination at both sites is followed by polyadenylation performed by poly(A) polymerases, TRAMP components Trf4/5 and the major mRNA

enzyme Pap1 (Fig. 6C) (see 1.3.3.3). Similarly to protein coding genes, coupling of sn/snoRNA termination and 3' end processing also involves polyadenylation, though the role of snoRNA poly(A) tail, unlike that of mRNA, is to facilitate their 3' end processing or degradation by the nuclear exosome/Rrp6 (251, 254, 255).

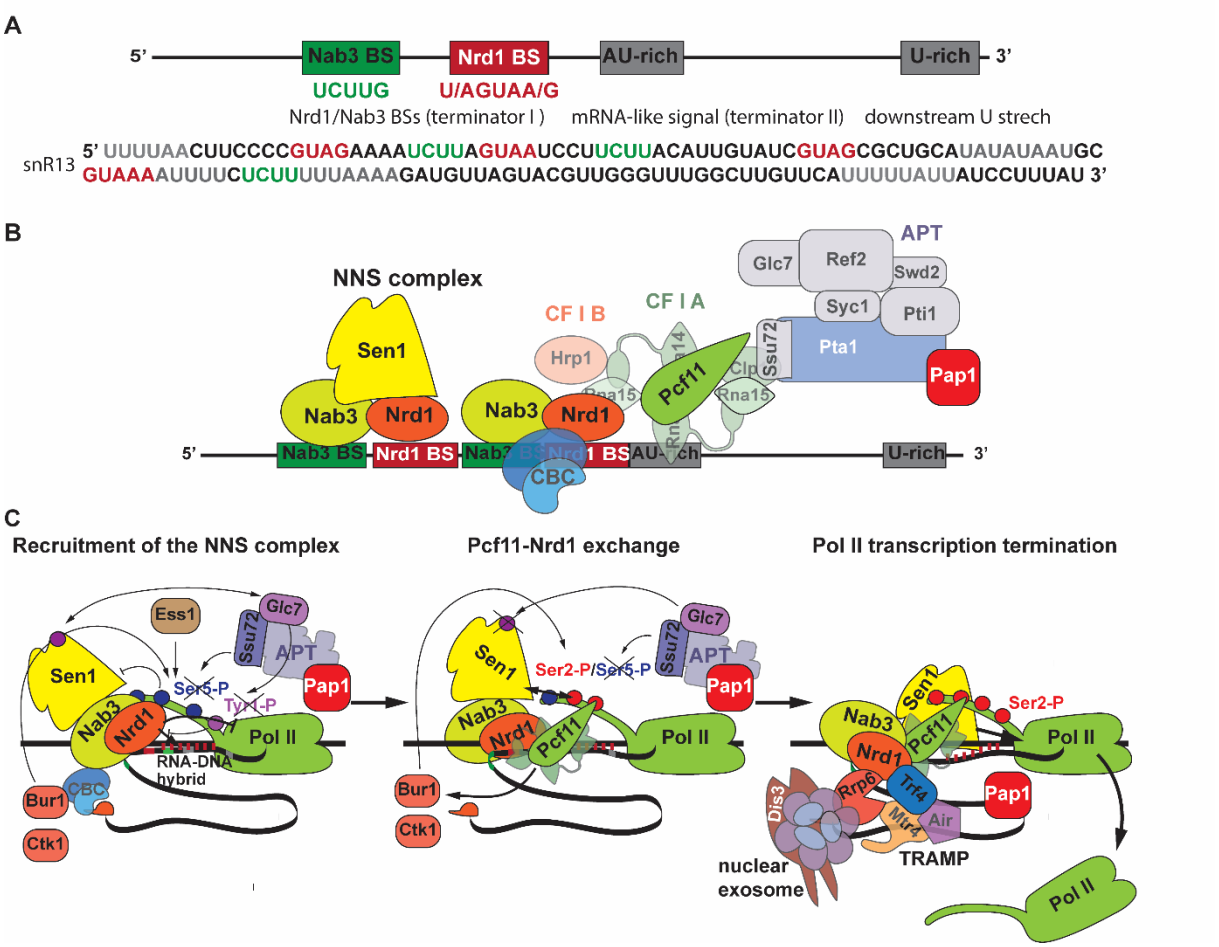

Figure 6. Polymerase II transcription termination at noncoding genes. *(**A**) NNS-dependent terminators: Nrd1 (red) and Nab3 (green) binding sites (BS), AU-rich elements resembling mRNA terminator signals and downstream U stretch (grey), according to* (206, 220, 252). *Pre-snR13 3' end sequence is shown below to exemplify the spacing and composition of these motifs in the RNA sequence. (**B**) Suggested organisation of the NNS-CFI-APT complex on noncoding RNAs, according to* (206). *Sub-complexes are colour-coded as in Figure 5. The NRD-CFI-APT complex has not been biochemically purified, but inferred from genetic studies. (**C**) Current model of NNS-dependent Pol II transcription termination, based on* (41, 207). *Lines with arrowheads indicate either activation, catalysed reaction or interaction (arrow heads at both ends), while lines with barheads indicate inhibition. Circles on Pol II CTD tail indicate phosphorylated residues and are colour-coded as on Figure 1. Purple circle on Sen1 represents phosphorylated residue dephosphorylated by Glc7. RNA is drawn as a black ribbon, terminator elements and termination factors are coloured as on schemes presented above. DNA template is drawn as a black line. Red bars indicate interaction between RNA and DNA. Description in the text.*

Initially, NNS-dependent termination was demonstrated mainly for sn/snoRNAs but more recent genome-wide and high-resolution studies have expanded the list of NNS targets that include both stable and unstable ncRNAs and some mRNAs. NNS role is therefore not only restricted to Pol II ncRNA termination but seems to be extended to widespread regulation of

gene expression either directly or by degradation of intergenic and antisense ncRNAs (248). It appears that 80% of yeast sn/snoRNAs biogenesis relies upon the NNS complex, as revealed by RNA-seq studies upon Nrd1 conditional depletion from the nucleus (256). In addition, noncoding Nrd1-unterminated transcripts, called NUTs, from nucleosome-depleted regions were shown to be common at promoters for divergent species or at most 3' regions of genes for antisense NUTs.

According to genome-wide studies of Pol II occupancy in *sen1* mutant or upon Nrd1 depletion, transcription termination of only a small subset of mRNAs is affected by NNS dysfunction and this concerns not only not only expression of short genes (~200–550 bp), but also these of the average length (~1.2 kb) (247, 256). Remarkably, binding of Nrd1 and Nab3 proteins was shown for more than 1,000 mRNA targets by cross-linking approaches (257, 258). In several cases NNS-dependent terminators are localized in 5' regions of many protein coding genes and the association of Nrd1/Nab3 as well as APT subcomplex with this regions probably promotes early termination coupled to degradation of non-functional transcripts, acting as a attenuation type of gene expression regulation. Based on genome-wide transcript profiling in *sen1* and *nrd1* mutants and recent high-throughput studies it is considered that this mechanism may concern even hundreds of mRNA genes (256, 258), including *NRD1* and other termination or RNA surveillance factors such as *PCF11, HRP1, MUD1* (U1 snRNA binding protein) and *TRF4*. Accordingly, expression of these proteins is upregulated when components of the NNS-dependent termination pathway are mutated (206, 207, 248). The best characterized example is expression of *NRD1*, controlled via a negative feedback loop involving NNS-dependent termination with a suboptimal efficiency that depends on cellular levels of Nrd1 protein (259). Similar mechanism is employed by the ribosomal protein Rpl9 that autoregulates its own synthesis, probably by masking CPF–CF termination signals and enabling NNS-dependent termination and nuclear degradation of abortive transcripts (260).

NNS-dependent termination can also operate as the fail-safe mechanism in situations when Pol II reads through the normal terminator, especially for highly expressed genes. Often this termination is coupled to degradation of the nascent transcript by the exosome, unless it is prevented by mRNP assembly factors or readenylation by Pap1 (206). Nrd1/Nab3 binding sites downstream of the PAS, unlike these located within gene body, are recognized as terminators due to changes in Pol II CTD phosphorylation status and other rearrangements of the elongation complex (142).

Transcription termination of CUTs is also dependent on the NNS-complex, although it is connected with transcript degradation instead of its further processing (206, 207, 248). This is probably caused by the lack of protection provided by specific RNP proteins that might otherwise stabilize these transcripts. Following transcription termination, CUTs are oligoadenylated by the TRAMP complex and then degraded by the exosome associated with Nrd1 and TRAMP. Exosomal degradation of these RNAs is so rapid that they are only detected when the exosome is not fully functional (261, 262). NNS involvement in CUTs transcription termination and degradation has been shown in many genome-wide studies as well as by directed experiments for some well-characterized genes such as *PHO84* (248). RNA-seq analyses of Nrd1 depletion effects revealed widespread function of Nrd1 in terminating intergenic and antisense transcripts and controlling their level and length (256, 263). The regulation of CUTs termination is important for proper transcription of functional transcription units, as they are widespread, often originate from bidirectional promoters associated with coding genes and sometimes overlap with mRNA genes in sense or antisense orientation. This bidirectionality, implicated in regulation of gene expression, arises from two adjacent PICs that usually give rise to a functional and antisense transcript (248).

A similar regulation strategy has been observed in humans, where PROMPTs (promoter-proximal transcripts) are the equivalent of CUTs (207). Although transcription termination of these short RNAs involves the PAS and CP machinery, it also depends on factors analogous to these directing termination of snRNA and histone genes. The important players that stimulate this process are CBC and ARS2 (see 1.2.2), that probably recruit the CFII component CLP1 specifically to promoter proximal PASes (95, 96). Similarly to the yeast NNS complex, CBC–ARS2 associates with the NEXT complex, which promotes PROMPT degradation, important for the suppression of promoter bidirectionality in humans. It is not entirely clear however, whether this pathway interferes with the recently described Xrn2- and Dcp1/Dcp2-dependent “torpedo” termination at promoter proximal regions of human genes (192).

In contrast to yeast, most of mammalian snoRNAs are encoded within introns and their biogenesis in not coupled to transcription termination (109). However, mammalian snRNAs (U1, U2, U4, and U5) are encoded by independent transcription units and their 3’ end maturation depends on specific snRNA promoters, distinct from those of protein coding genes, and requires a 3’ box element located 9–19 nt downstream of their mature 3’ end (204, 264). A large Integrator complex is responsible for termination and 3’ end processing of for

mammalian snRNAs. It is recruited to the EC through Pol II CTD phosphorylated on Ser7 and Ser2 and one of its components (see 1.1.1), IntS11, belonging to the β-CASP family, is believed to perform endonucleolytic cleavage of the nascent transcript (22, 264, 265). In turn, APT and CF subunits function as transcription terminators for human snRNA genes with little, if any, role in their 3' end processing (266). Pol II release has been linked to the particular chromatin structure at snRNA genes and to the action of the negative elongation factor NELF, which is involved in promoter-proximal pausing at protein coding genes (266, 267). Another complex shown to be engaged in human snRNA transcription termination and the recognition of early 3' processing signals is CBCA (95, 96). Moreover, ARS2 stimulates PHAX association with CBC on snRNPs that couple transcription termination to nuclear export (96). The CBCA complex was proposed to be the functional analogue of the yeast NNS complex.

1.3.2.1 Nrd1

Nrd1 (nuclear pre-mRNA down-regulation) is an essential RNA-binding protein that governs termination of ncRNAs and some mRNAs, in association with another RNA binding protein Nab3 and helicase Sen1 (36, 103, 268, 269). Nrd1 consists of a central RNA recognition motif (RRM), N-terminal CTD interacting domain (CID), C-terminal glutamine/proline rich (Q/P-rich) domain and Nab3-interacting region. Importantly, Nrd1 CID, which is dispensable for yeast viability, preferentially binds Ser5-P CTD (268, 270–273).

Nrd1/Nab3 heterodimers cooperatively bind RNA in a sequence specific manner, it has been shown that Nrd1 affinity to its target is enhanced within the heterodimer (274). Nrd1 and Nab3 binding motifs in snoRNA terminators, GUAA/G and UCUU, respectively, are located downstream from the mature sn/snoRNA 3' end and their number vary from one to more than ten (Fig. 6A) (244, 259, 274–277). Direct Nrd1 and Nab3 interaction with RNA was confirmed and their binding sites within ncRNA terminators precisely mapped using genome-wide cross-linking methods such as PAR-CLIP (photoactivatable ribonucleoside-enhanced cross-linking and immunoprecipitation) and CRAC (cross-linking and analysis of complementary DNAs) (256–258, 278, 279). Consensus Nrd1 and Nab3 binding motifs from these studies closely match those obtained using the *in vivo* SELEX (systematic evolution of ligands by exponential enrichment) approach and mutagenesis of specific snoRNA terminators, but they were placed within longer supermotifs overrepresented in ncRNA terminators: U/AGUAA/G for Nrd1 and UCUUG for Nab3 (217, 244, 252, 275). Nrd1 and

Nab3 binding motifs are depleted from coding regions, but are mainly present within AT-rich intragenic regions of either short or long genes, some promoter proximal regions, ORF antisense orientations and tRNA elements (252, 256–258). The NNS complex seems to be involved in the regulation of expression of wide range of transcripts, though only Nrd1 cooperative binding to Ser5-P CTD and RNA motifs containing also downstream AU-rich elements results in transcription termination (252, 271). Nrd1/Nab3 binding sites are also not recognized when located within gene bodies more than 1000 bases downstream of promoters (259, 271).

The exact role of Nrd1 in ncRNA transcription termination, where Sen1 seems to be responsible for Pol II release, is not fully understood. Together with Nab3 it acts as a specific sensor of ncRNA terminators, recognizing sequence motifs described above. Multiple Nrd1-Nab3 heterodimers bound to RNA create a ribonucleoprotein platform that recruits Sen1, which is underrepresented in the cell (248). Nrd1 seems to be important for Pol II pausing prior to termination, as its depletion eliminates pausing and enables Pol II to transcribe through the termination site (220). Simultaneous binding of the NNS complex to Pol II CTD and the nascent transcript contributes to efficient transcription termination and was proposed to stimulate the formation of the RNA:DNA hybrids by bridging the nascent transcript and the DNA template (41, 273). Furthermore, Nrd1 stimulates Pcf11 recruitment, which in turn promotes Nrd1 release from the CTD, making it accessible for interactions with different termination and 3’ end processing factors (41). This exchange of NNS interacting partners is proposed to coordinate transcription termination and nascent RNA degradation. Mechanism responsible for coupling ncRNA termination with 3’ end processing or degradation depends on Nrd1 interaction with a CTD mimic, so called Nrd1-interacting motif (NIM) of Trf4 (poly(A) polymerase from the TRAMP complex) (280). Interaction with Pol II CTD and Trf4 are mutually exclusive and Nrd1-Trf4 interaction is required for optimal exosome and TRAMP activity (280). Nrd1 also associates with the nuclear exosome complex through direct binding with Rrp6 (103, 280). This interaction is not mediated by CID, although replacing Nrd1 CID with that of Rtt103 reduces NNS binding to both Rrp6 and Trf4, while CID deletion completely abolishes association with the exosome (273, 280). Moreover, Nrd1 interacts with CBC, although the role of this interaction and its importance for ncRNA transcription termination and 3’ end processing is not clear. The current model of NNS-dependent transcription termination is presented in Fig. 6C.

The NNS complex activity is modulated in response to external factors. Although it is not fully understood how it is achieved, one possibility could include Nrd1 post-translational modifications that are influenced by the level of Nab3 (206). It turns out that Nrd1, as well as other two components of the NNS complex, is phosphorylated, but no particular kinase has been attributed to carry out these modifications (270, 281, 282). There is evidence that the Ras signalling pathway negatively affects NNS-dependent termination, however, it is not connected to Nrd1 phosphorylation, but may be associated with the transcriptional regulatory mechanism involving interaction of Nrd1/Nab3, Pol II transcription factors Spt4/5 and Ras with the CTD within the early EC (283). In turn, NNS-dependent termination plays an important role in the regulation of the cellular response to nutrient availability. It causes rapid suppression of some genes upon glucose depletion that is associated with Nrd1 dephosphorylation and sequestration of Nrd1/Nab3 in novel nuclear speckles (283). Nrd1 was also implicated in controlling cell size and mitochondrial abundance and monitoring the expression of the cell cycle regulator *CLN3*.

The N-terminal 163 amino acids of Nrd1 are essential for viability and deletions of Nrd1 RRM, P1Q and RE/RS domains or mutations in RRM or CID domains confer temperature sensitivity (268, 270). Nrd1 mutations or its nuclear depletion cause defective termination of many NNS-dependent transcripts resulting in accumulation of over 1,500 extended ncRNAs referred to as NUTs (228, 256). Generation of NUTs leads to severe changes in gene expression, predominately through transcription interference.

1.3.2.2 Nab3

Nab3 (nuclear polyadenylated RNA-binding) was identified as a nuclear RNA-binding protein related to vertebrate hnRNP-C through its RRM (284). Besides the conserved RRM that is essential for yeast viability it is composed of low complexity domains: non-essential N-terminal, aspartic/glutamic acid rich region of unknown function and essential C-terminal glutamine/proline rich region (274, 285, 286). The latter is similar to prions and can polymerize into an amyloid-like structure and form hydrogel *in vitro* (285, 286).

Nab3 interacts with Nrd1 forming a stable heterodimer through Nrd1 interacting domain (270, 274). It can also self-associate into a tetrameric form through a region of structural homology shared with hnRNP-C (287, 288). This tetramerization is independent of interactions with RNA and Pol II, but it is important for transcription termination, providing an additional tether between Nrd1/Nab3 heterodimers bound to multiple sequence motifs on the nascent RNA.

Nab3 recognizes UCUU motifs within ncRNAs terminators (244, 274, 275, 289). The extended, UCUUG supermotif is highly enriched in CUT terminators, where Nrd1 binding motif is often absent, so in contrast to snoRNAs, CUT terminators seem to be predominately recognized by Nab3 (219).

1.3.2.3 Sen1

Sen1 belongs to a conserved RNA/DNA helicase superfamily I and its helicase activity has been confirmed for the *S. pombe* orthologue (268, 290, 291). *S. cerevisiae* Sen1 consists of a leucine zipper motif, consensus elements for nucleoside triphosphate binding, and the nuclear localization signal (290). Its large surface is responsible for interactions with many factors, proposed to play an important role in coordinating NNS-dependent processing and termination (292). In contrast to abundant Nrd1 and Nab3 proteins (5,800 and 19,600 molecules per cell, respectively), only about 125 molecules of Sen1 are present in a cell (293). Sen1 binds Pol II CTD preferentially in the Ser2-P state and this interaction is important for ncRNA termination *in vivo* as it promotes Sen1 function, although it is dispensable for Sen1 termination activity *in vitro* (37, 292, 294–297).

Sen1 interacts with the protein phophatase-1 catalytic subunit Glc7 that, as a CPF component, dephosphorylates Tyr1-P CTD and is required for efficient termination at both sn/snoRNA and protein coding genes (46, 298, 299). The interaction between Sen1 and Glc7 can facilitate recruitment of Sen1 to the APT complex and of Glc7 to the Nrd1/Nab3 complex (Fig. 6C). Considering that Glc7 was shown to dephosphorylate Sen1 *in vitro*, these associations may enable Glc7 to dephosphorylate Sen1 and possibly also Nrd1.

The large network of Sen1-binding partners include all snoRNP core proteins, endonuclease Rnt1 and deoxyribonuclease Rad2 (36, 292, 298). Sen1 interaction with Rnt1 seems to play a role in exonucleolytic processing of Rnt1 cleavage products (292). In turn, association with Rad2 suggests Sen1 involvement in DNA repair, that was demonstrated for its mammalian homologue, Senataxin (300–302).

As the only catalytic subunit of the NNS complex Sen1 is proposed to be responsible for Pol II release from the DNA template. Although the exact mechanism is not completely understood, it probably involves ATP hydrolysis by Sen1 triggering EC dissociation in a similar manner to the bacterial termination factor Rho (297). *In vitro* Sen1 is sufficient to elicit termination independently of Nrd1 and Nab3 and probably in a sequence independent manner. This mechanism involves both ATPase and helicase activities of Sen1, as well as its interaction with the nascent transcript (minimum 15 nt long). In turn, NNS-dependent

transcription termination *in vivo* is enabled by a temporal and spatial window, established by the kinetic competition between elongating Pol II and Sen1 helicase action (303). Together, these data support a termination model wherein Sen1 translocates along the nascent RNA to catch up with Pol II and trigger its dissociation from the template, similarly to the bacterial termination factor Rho, while Nrd1 and Nab3 act mainly upstream, ensuring efficient and specific recruitment of Sen1 (207, 297).

On the other hand, Sen1 was proposed to function in dismantling of RNA-DNA hybrids (R-loops) that form between the nascent transcript and the unwound template (304). Such role of Sen1 is supported by genome-wide cross-linking studies that showed its association with many transcribed genes, also independently of the Nrd1/Nab3 heterodimer (37, 278, 279, 305). R-loops may play a role in termination by slowing down Pol II and providing time for association of termination factors such as Pcf11. However, this alternative appears incompatible with the *in vitro* data claiming that R-loops cannot be removed by Sen1 during transcription and they actually inhibit termination *in vitro* (297). It has been proposed that Sen1 instead of actively removing RNA-DNA hybrids may suppress their formation by association with the nascent RNA. Moreover, Sen1 has been shown to prevent transcription-associated recombination, genome instability and replication fork defects, possibly via its ability to control accumulation of R-loops (305, 306). In accordance, human Senataxin (SETX) has been proposed to facilitate Pol II termination by resolving RNA-DNA hybrids to allow access of Xrn2 (307). Senataxin mutations cause pleiotropic transcription processing and termination defects that are manifested in cerebellum and result in neurodegenerative disorders such as AOAII (ataxia ocular apraxia type II) and ALS4 (amyelotropic lateral sclerosis 4) (307–309). Although Senataxin does not rescue *SEN1* mutations in yeast, some of the disease-related SETX mutations cause termination defects when introduced into yeast *SEN1* (296).

Yeast Sen1 is essential for viability, however only the C-terminal 1,214 amino acids are indispensable, whereas the N-terminal 898 amino acids are nonessential (290). *SEN1* mutations or lack of Sen1 helicase activity cause most severe termination defects not only for Pol II genes, including most snoRNAs, but also for rDNA transcribed by Pol I (206, 228, 247). Defective termination in *sen1* mutants is modulated by mutations in *RPB1*, respectively enhanced by faster polymerase and suppressed by its lower processivity (303). Hence, kinetic competition between elongation and Sen1 activity was proposed to create a spatial and temporal window, mainly important for termination of short Pol II transcripts.

In addition to direct termination phenotype, *sen1* temperature sensitive mutants or Sen1 depletion exhibit pleiotropic effects, including defects in pre-tRNA splicing and rRNA and snRNA processing, as well as complications in nuclear fusion and chromosomal stability (36, 310–312). It is possible that these dysfunctions are indirectly caused by transcriptional readthrough into relevant genes, as is the case for tRNA splicing ribonuclease *SEN2* (269).

1.3.2.4 The CP machinery and additional factors

Several studies have shown that many components of the CP machinery contribute to ncRNA transcription termination, but their exact role is often not clear (Fig. 6B). In particular, mutations of the CF IA subunits, Rna15, Rna15, Clp1 and Pcf11, cause transcriptional readthrough at several sn/snoRNA genes, while inactivation of Rna14 and Rna15 leads to delocalisation of polyadenylated snoRNPs from the nucleolus to discrete nucleoplasmic foci (228, 250, 313). Remarkably, cleavage is not required for yeast sn/snoRNA termination and depletion of Ysh1 (CPF) or *PCF11* mutation that blocks cleavage does not affect this process (228). Although a cold-sensitive *ysh1* mutant showed defective termination at some snoRNA genes, no genome-wide snoRNA termination defect was observed upon Ysh1 depletion from the nucleus (212, 220). In contrast, *PCF11* mutations disabling its binding to CTD impair Pol II release and snoRNA termination, probably due to ineffective competition with Ser5-P bound Nrd1 (51, 228) or by preventing Pcf11-Nrd1 CP-independent cooperation in ncRNA termination (41). Pcf11 is recruited downstream of Nrd1-binding, over distal NRD terminator regions, and facilitates Nrd1 release from CTD. On the other hand, Pcf11 binding to CTD enhances Ser2 phosphorylation that in turn promotes Sen1-CTD interaction that is crucial for its termination activity *in vivo* (Fig. 6C) (37, 292).

Core CPF subunits are not necessary for sn/snoRNA transcription termination, whereas all APT components, namely Glc7, Swd2, Pti1, Pta1, Ref2, Ssu72, were shown to be required (Fig. 6B) (16, 211, 217, 228, 298, 314, 315). Pti1 was proposed to uncouple cleavage from polyadenylation, as its overexpression inhibits polyadenylation (314). This control of snoRNA poly(A) tail length may serve to ensure their nuclear localisation, as pre-mRNA polyadenylation is directly coupled to export (206). Both Pti1 and Ref2 are important for snoRNA termination and act as multicopy suppressors of *pcf11* mutations, but are not essential for pre-mRNA processing *in vitro*. In addition, Ref2 interacts with the snoRNP specific protein Nop1 (316). Similarly, the N-terminal region of Pta1 is necessary for snoRNA termination being dispensable for pre-mRNA 3’ end processing (16), whereas its conserved C terminus, that connects it with the core CPF, is essential for mRNA cleavage and

polyadenylation (16, 211). In turn, deletion of the N-terminal region of Pta1, a CPF scaffold protein interacting with Ssu72, Pti1 (APT), and Ysh1 (core CPF), weakens its association with the CTD phosphatase Ssu72, affecting its stability, CTD Ser5-P dephosphorylation, snoRNA transcription termination and gene looping (16). Ssu72 CTD phosphatase that removes Ser5-P and Ser7-P marks at terminators, plays an important role in both CP- and NNS-dependent transcription termination (14, 217, 315, 317). Ssu72 inactivation causes transcriptional readthrough and impairs reinitiation, as it is crucial for CTD dephosphorylation prior to Pol II engagement into new PIC (14). Another APT subunit, Swd2, is also a component of the Set1-COMPASS complex, which modifies lysine 4 of histone H3 (H3K4) (318, 319). These two roles of Swd2 are probably separable. Depletion of Swd2 results in loss of APT recruitment, leading to snoRNA termination defect. This phenotype can be overcome by *SET1* deletion, or overexpression of Ref2 or non-functional C-terminal fragment of Sen1 (298, 318). Recent data suggests that histone modification by the Set1 complex can be antagonistic to NNS-dependent termination (320). It is possible that Swd2 recognizes the early EC, perhaps via sensing CTD-Ser5-P, and coordinates the simultaneous or sequential recruitment of both complexes. Consistently, *SET1* deletion increases CTD-Ser5-P level and affects binding of NNS and APT complexes to snoRNA terminators (320). Similar termination defects at snoRNA genes is caused by Glc7 downregulation or its dissociation from CPF in the absence of Ref2 or Swd2 (298). Glc7 is a protein phosphatase that directly interacts with Sen1 and dephosphorylates this protein *in vitro*. Loss of Glc7 from CPF can be compensated by its direct recruitment to snoRNA terminators via interaction with Sen1.

Similarly to protein coding genes, Rpb3/11 heterodimer as well as the antiterminator Paf1 complex are involved in snoRNA transcription termination. The latter interacts with Pol II and mediates histone modifications during elongation and was proposed to facilitate Nrd1 recruitment to EC (68). Also co-transcriptional association of specific pre-sn/snoRNP proteins that stabilize RNA three-dimensional structure ensuring its stability is important for proper transcription termination and 3’ end processing (316, 321).

Recent findings suggest that exonuclease Rrp6 associated with the nuclear exosome is not only involved in processing/degradation of transcripts terminated via the NNS-dependent pathway but is also engaged in the termination process itself (Fig. 6C) (322). Genome-wide studies revealed that Rrp6 deletion leads to decreased expression of the majority of sn/snoRNAs and causes changes in Pol II occupancy at some but not all NNS-dependent

transcripts. For instance, transcription termination of snR71, snR4, snR34, CUT281, CUT882 as well as *NRD1*, *HRP1* and *YPL222C-A* was shown to be Rrp6 dependent, while snR13 does not require Rrp6 for proper termination (322–324). Rrp6 directly interacts with Nrd1 and this interaction has been postulated to play an important role in termination of a subset of specific NNS-dependent RNAs (103, 273, 280). Thus, Rrp6 can regulate both degradation and the efficiency of NNS-dependent termination of sn/snoRNAs, CUTs, NUTs as well as some SUTs and mRNAs and is therefore implicated in gene expression regulation through transcription interference. There are also some indications of the involvement of the core exosome nuclease Dis3 in NNS-dependent termination of snR4 and snR34 (323). In turn, *S. pombe* exosome was recently suggested to be directly engaged in transcription termination according to a novel, provocative, double-torpedo model, based on observations that depletion of Dis3 or Rrp41 (but not Rrp6) induces a termination defect in approximately 30% of Pol II transcripts (325). This model proposes that the exosome can promote Pol II release by degrading the 3' extension in backtracked Pol II, possibly complementing the concurrent 5'-3' torpedo degradation by Rat1 (325). However, it is still debatable whether the 3' RNA extension presented by the backtracked Pol II is a substrate sufficiently long for the exosome.

1.3.3 Endo- and exonucleolytic 3' end processing

Transcription termination at sn/snoRNA genes generates 3' end extended precursors (264). Thus, in contrast to the CP pathway, NNS-dependent termination is directly coupled to TRAMP/exosome-dependent processing of nascent transcripts. In the case of CUTs and many other unstable ncRNAs TRAMP/exosome action leads to total transcript degradation. It was proposed that Nrd1 is important for determining the fate of NNS-dependent transcripts between their processing or degradation (103). In addition, this decision is most likely dictated by sn/snoRNA-associated proteins that serve as a barrier against further exonucleolytic trimming and in this way define mature RNA ends. Some yeast pre-snoRNA and all pre-snRNAs contain a stem-loop structure located downstream of their mature 3' ends, which is recognized and cleaved by Rnt1, generating the entry site for further TRAMP/exosome action (143, 326–329).

1.3.3.1 Rnt1

Rnt1 endonuclease and its role in *S. cerevisiae* RNA metabolism are described in chapter 1.2.4.1. The 3' end extensions of pre-snRNAs and some pre-snoRNAs (e.g. U3) are cleaved by Rnt1 prior to exonucleolytic trimming. One of the best-characterized example is

maturation of U5 snRNA that exists in two forms, U5L and U5S, with different 3' ends generated by Rnt1cleavage at two sites (328). The cleavage choice is dictated by interaction of Rnt1 with Sen1, important for efficient maturation of U5L (36). The role of Rnt1 in 3' end processing of other sn/snoRNAs is described in chapter 1.5.

Rnt1 is also involved in the alternative fail-safe mechanism of transcription termination when PAS is not recognized by a standard termination machinery (141, 142). In this case, Rnt1 cleavage of a steam-loop structure generates an entry site for the Rat1 exonuclease promoting torpedo termination that is usually not coupled to polyadenylation. The resulting transcripts that are not stabilized by the poly(A) tail undergo rapid degradation by the TRAMP/exosome complex (142). However, if Rnt1 cleavage occurs close to the weak pA site, 3' end processing machinery is recruited and the transcript is polyadenylated by Pap1 (141). Association of Rnt1 with the termination site probably occurs via the interaction with the Pol II CTD. Rnt1 was shown to interact with the CTD mutated in Ser2 or Ser5 positions, required for termination by the two major mechanisms (141). Rnt1 cleavage provides additional possibility to terminate transcription and remove aberrant transcripts to avoid transcription interference.

Out of several cases of Rnt1-dependent Pol II transcription termination, the best-characterized example is *NPL3*, with an inefficient PAS-dependent mechanism (141, 330). Binding of Npl3 to its own PAS autoregulates its usage by causing transcriptional readthrough. This effect is further enhanced by Sen1, possibly by influencing the structure of the *NPL3* transcript (247). These readthrough transcripts become cleaved by Rnt1 and are degraded by TRAMP/exosome complex (141, 142).

1.3.3.2 Nuclear exosome

The exosome complex is the main eukaryotic 3'-5' exonuclease that is a key player in numerous RNA decay and processing pathways (331, 332). Two complexes of slightly different composition are present in the cytoplasm and the nucleus where they participate mainly in mRNA turnover and maturation of stable RNA species, respectively. Furthermore, both exosome complexes take part in RNA surveillance pathways that eliminate aberrant RNAs, while the nuclear exosome is engaged in degradation of RNA processing byproducts and ncRNAs such as asRNAs or CUTs.

The eukaryotic exosome consists of nine subunits that form a ring-like structure (referred to as Exo-9) (331–333). A central channel is composed of bacterial RNase PH homologues, Rrp41- Rrp46 and Mtr3. Three remaining subunits, Rrp4, Rrp40 and Csl4, that contain S1

and/or KH-type RNA binding domains, form a cap on the top of the channel. Despite of the structural similarity to the archaeal exosome-like complex and the bacterial polyribonucleotide phosphorylase (PNPase), eukaryotic Exo-9 complex, possibly with the exception of RRP41 subunit in some plant species, does not display catalytic activity (334–336). Instead, it serves as a scaffold for interactions with outer catalytic subunits and additional factors (335, 337). Its additional important role is to bind the RNA substrate and 'present' it for catalytic subunits.

The major catalytic exosome subunit is Dis3 (also referred to as Rrp44), which in yeast is associated with the Exo-9 ring of both the nuclear and cytoplasmic complexes, forming the Exo-10 complex (335, 338). In human cells, two Dis3 homologues, predominantly nuclear DIS3 and strictly cytoplasmic DIS3L, form the respective Exo-10 complexes (339). The 110 kDa Dis3 is composed of five domains: an N-terminal Pilus-forming N-terminus (PIN) domain, two cold-shock domains (CSD1 and CSD2), a central ribonuclease domain (RNB), homologous to bacterial RNase II/R, and a C-terminal S1 domain (332). It displays 3'-5' processive hydrolytic exoribonuclease activity, conferred by the RNB domain, and endonuclease activity located in the PIN domain (340–342). Surprisingly, of two human variants only nuclear DIS3 has endonucleolytic activity (339). Additionally, PIN and CSD1 domains interact with Rrp41/Rrp45 and Rrp43 subunits of the core Exo-9 ring, respectively. According to recent biochemical and structural data, Dis3 is located at the bottom, on the opposite side of the trimetric cap (337, 342–345).

Additional catalytic cofactor, Rrp6, homologue of bacterial RNase D, interacts with yeast nuclear exosome, forming the Exo-11 complex (334, 338). Human RRP6 is present in the nucleus and probably also nucleolus, where it associates with the Exo-10 or Exo-9 (lacking DIS3), respectively (339). Rrp6 is composed of five domains: an N-terminal PMC2NT domain, an EXO domain with a distributive 3′-5′ exoribonuclease activity, a HRDC domain (helicase and RNase D C-terminal domain), a predicted HRDC2 domain, and a C-terminal domain (CTD) (332). Rrp6 is located on the other side of the channel, in the vicinity of the trimeric cap, making contacts with Csl4, Mtr4 and Rrp43 (333, 345, 346). The C-terminal region of Rrp6 is predicted to stabilize RNA binding by the core exosome, extending it through the channel (345).

Before reaching Dis3 RNB domain, RNA substrate is threaded through the channel to obtain at least 31-33 nt long single-stranded 3' end (334, 335, 337, 347). RNA unwinding, probably assisted by cap proteins, occurs by its passage through the Exo-9 barrel in a single-

stranded conformation, while interacting with channel subunits in a sequence-unspecific manner (345). Threading appears to protect the substrate from Dis3 endonuclease active site, which faces the solvent (337, 345). Both catalytic activities of Dis3 probably cooperate on most substrates, particularly those containing strong secondary structure when endonucleolytic cleavages facilitate exonucleolytic degradation (333, 340, 348). In turn, the upper portion of the channel, comprising the S1/KH ring, is important for Rrp6 activity (349). Therefore, shown *in vitro* and *in vivo*, threading through the core exosome channel appears to influence all three exosome RNase activities (337, 349, 350). Furthermore, Dis3 and Rrp6 seem to cooperate, as Rrp6, independently of its catalytic activity, stimulates both endo- and exoribonucleolytic action of Dis3, while Dis3 lacking exonuclease activity severely inhibits Rrp6 activity (349). Notably, highly structured Dis3-dependent substrates are often associated with both catalytic subunits, but not the core exosome. The alternative recruitment pathway to the Dis3 catalytic site was proposed to be aided by docking to Rrp6 and other exosome cofactors (348). It probably allows the direct access of shorter single-stranded RNA regions to Dis3 or Rrp6.

The choice of degradation pathways of particular substrates can be additionally regulated by cofactors competing for interactions with the exosome core or by influencing activities of Rrp6 and/or Dis3 (332, 333). In *S. cerevisiae* two related DExH-box RNA helicases, Mtr4 and Ski2, activate the exosome in the nucleus and cytoplasm, respectively. Although Mtr4 can function alone in assisting RNA unwinding, it also acts in the context of the TRAMP complex (Trf4/5–Air1/2–Mtr4 polyadenylation complex) which, in co-operation with the NNS complex interacting with Rrp6, recruits the exosome to degrade or process certain Pol II ncRNAs (see 1.3.3.3) (103, 142, 351). Similarly, in humans, the equivalent NEXT complex, comprising hMTR4, the Zn-knuckle protein ZCCHC8, and the putative RNA binding protein RBM7, binds the exosome to promote degradation of PROMPTs (352). The nuclear exosome interacts also with Rrp47/C1D via Rrp6 and Mpp6/MPP6 to process structured RNAs such as precursors of 5.8S rRNA and sn/snoRNAs (332, 353). The latter is mediated through the interaction between Rrp47 C-terminal region and box C/D small nucleolar ribonucleoproteins Nop56 and Nop58. Rrp47 is important for the recognition of Rrp6 substrates prior to their degradation, and was proposed to promote NNS-dependent transcription termination in concert with Rrp6 and Trf4 (251, 255, 259, 324). However, Rrp47 and Rrp6 also have separate functions in Rrp6-mediated RNA surveillance and processing. For example, Rrp6

can be at least partially replaced by another exoribonuclease, Rex1, in snoRNA maturation (255).

Exosome is responsible for degradation and/or processing of bulk transcripts, which reflects tight interconnection between 3' end processing and RNA surveillance pathways (332, 333). In yeast it is engaged in maturation of 3' ends of many RNA species, including 5.8S rRNA, sn/snoRNAs and even some mRNAs (333). In the case of *NAB2* and *CTH2* mRNAs, their processing from 3'-extended precursors mediated by the NNS and nuclear exosome/TRAMP complexes may ensure additional regulation of their expression (258, 333, 354, 355). Notably, the levels of *NAB2* are controlled by a competition between Rrp6-dependent surveillance and polyadenylation. Specifically, Nab2 binding to its own mRNA promotes its degradation by Rrp6, whereas displacement of Nab2 by Rrp6 may contribute to polyadenylation efficiency and association with poly(A) binding proteins (355).

Another important role of the nuclear exosome is to eliminate many aberrant or superfluous transcripts and processing intermediates, both coding and noncoding. These substrates include excised pre-rRNA spacers, most misassembled RNPs and misprocessed precursors, for example unspliced or 3'-unprocessed pre-mRNAs and mRNAs trapped in the nucleus that would otherwise accumulate in various RNA processing and export mutants (100, 356–359). Multiple high-throughput RNA-seq, CRAC and PAR-CLIP analyses identified additional classes of Dis3/DIS3 and Rrp6/RRP6 targets, including CUTs and SUTs, PROMPTs (in humans), snoRNAs, and, most prominently, pre-tRNAs and other Pol III transcripts (257, 348, 360). In addition, these approaches revealed that Dis3 and Rrp6 have both overlapping and specific functions. Although Rrp6 normally acts within the context of EXO-11, it can also degrade some targets independently of the exosome, as subset of its functions are not affected by loss of Dis3 or any other Exo-9 components (333, 361). The nuclear exosome has been also proposed to be engaged in transcription termination, which is described in chapter 1.3.2.4 (322–325).

The core exosome subunits and Dis3 are all strictly essential for yeast viability, whereas Rrp6 is not essential, but its deletion causes a temperature-sensitive growth phenotype. Exosome function has been recently connected to several other cellular pathways, including an role in generating antibody diversity (362). It was also linked to human diseases such as polymyositis-scleroderma overlap syndrome and multiple myeloma (363, 364).

1.3.3.3 TRAMP

The role of the TRAMP (Trf4/5–Air1/2–Mtr4 polyadenylation complex) complex in the 3' end processing or degradation is to activate and modulate catalytic activity of the nuclear exosome (365). TRAMP oligoadenylates transcripts that are further digested by Rrp6 and Dis3, assisting in transcript recognition and enhancing exonuclease activity. There are two TRAMP complexes present in the *S. cerevisiae* nucleus, namely TRAMP4 and TRAMP5 that contain Trf4 or Trf5, respectively. Trf4 and Trf5 are non-canonical poly(A) polymerases that lack RNA binding domain and therefore require one of two RNA binding proteins, Air1 or Air2, for substrate recruitment (366, 367). Although Trf4 is able to add quite long poly(A) tails to its substrates *in vitro*, recent *in vivo* studies showed that its catalytic activity is modulated by Mtr4 suppressing polymerization after addition of 3–5 adenosine residues and that distribution of oligo(A) tails added by TRAMP peaks around 4–5 adenosines (257, 368). Such a short poly(A) tail is not recognized by poly(A) binding proteins, Pab1 and Nab2, that require 12 or 20 nucleotide overhangs, respectively, which prevents transcripts polyadenylated by TRAMP from being exported to the cytoplasm (368, 369).

Air1/2 proteins, having 5 zinc knuckle RNA binding motifs, contribute to substrate specificity of both TRAMP complexes (365, 367, 370, 371). TRAMP4 usually contains Air2, while TRAMP5 only contains Air1 (261, 366, 372, 373). Air1 and Air2 have non-redundant functions in polyadenylation of different mRNA sets, while Air2 additionally has a strong preference for snoRNAs (374). Consistently, favoured TRAMP4 and TRAMP5 substrates complexes also differ. Trf4 is required for 3' end formation of *CTH2* and polyadenylation of rRNAs and pre-sn/snoRNAs, while Trf5 enhances degradation of aberrant rRNAs, *NAB2* and noncoding transcripts from intergenic spacer region (354, 355, 373, 375, 376).

The mechanism of substrate recognition by Air proteins is not fully understood. It was proposed that zinc knuckles bind to specific and distinct (but not yet identified) sequence motifs in target RNAs or may recognize particular structures, e.g. irregular folds common within aberrant transcripts, and in this way recruit or exclude some specific RNA binding proteins (RBPs) (365). Such structures may either directly interact with Air proteins (or their associated factors) or prompt other factors dissociation from RNA 3' end, making it available for Air1 or Air2 binding. Other proteins associated with Air1/2 may also contribute to their substrate specificity. For instance, the Air2-containing TRAMP complex interacts with the NNS complex, but this interaction is mediated by binding of Nrd1 CID with the NIM motif of Trf4 (103, 280, 377). In turn, overexpressed Air1 copurifies with Hrb1, a poly(A) binding

protein involved in mRNA export, methylated by Hmt1 (365). Also differential localization may influence substrate specificity of TRAMP4/5 complexes, which are both localized to the nucleus, but Trf5 and Air1 seem to be slightly enriched in the nucleolus, in agreement with their preferential function in polyadenylation of rRNA precursors (371, 373, 378). However, Trf4 is able to shift from nucleolus to nucleoplasm upon nucleolar accumulation of rRNAs, suggesting that differential localization of TRAMP complexes is more likely a result rather than a cause of substrate specificity (379).

Both TRAMP complexes also comprise the nuclear 3'-5' RNA helicase Mtr4, that is composed of a central ATPase core, typical of DExH helicases, and a unique arch domain, similar to the KOW domain of ribosomal proteins and required for 5.8S rRNA processing and $tRNA_i^{Met}$ binding (380, 381). Mtr4 interaction with Trf4 and Trf5 is most likely direct and independent of Air proteins (372). Mtr4 binds poly(A) RNA and probably functions downstream of TRAMP-mediated oligoadenylation that facilitates RNA unwinding (365). Since Trf4 and Mtr4 are present in excess over Air1, Air2 and Trf5, it was proposed that they may also perform some TRAMP-independent functions (327, 372, 382, 383). For example, Mtr4 can act outside the TRAMP complex as an exosome cofactor in the processing of structured RNAs such as 5.8S rRNA, and a group of substrates may require TRAMP, but not Mtr4, to stimulate degradation by the exosome (351, 366, 372). It has been proposed that Mtr4 in the TRAMP complex is located in the vicinity of Rrp6 and the RNA entry site, on top of the exosome (345, 380, 381). Both complexes are bridged by the arch/KOW domain of Mtr4 that hands off the substrate to the exosome or Rrp6. Structural data suggest that it may form a pore through which target RNA is likely to pass before reaching the exosome core channel (380, 381). The precise mechanism remains unclear, but it was suggested that RNA-bound Mtr4 has to dissociate from Air-Trf to associate with the exosome.

The exact mode of exosome activation by the TRAMP complex is not fully understood. A single-stranded poly(A) tail seems to make the highly structured RNAs more favourable substrates for the sterically restricted exosome. Although Trf4 catalytic activity is required to stimulate degradation of unmodified pre-$tRNA_i^{Met}$, multiple studies have shown it is not necessary for activation and/or recruitment of the exosome (261, 351, 366, 372, 375, 383, 384). On the other hand, the 3'-5' helicase activity of Mtr4 is required for degradation of some tRNAs, likely producing a single stranded fragment that can be degraded by the exosome (366, 385). Mtr4 preferentially binds to tracts of adenosines and unwinds substrates containing overhangs of 5–6 nucleotides (365). Its activity seems to be stimulated by

Trf4/Air2 by increasing its ATP affinity and rate of strand separation (386). However, transcripts that naturally contain 3' oligo(A) tail long enough to recruit Mtr4 and/or the exosome are not necessarily polyadenylated by Trf4. Another factor influencing RNA fate is the kinetics of its maturation (365). Defects within transcripts that stall the processing machinery may lead to its dissociation, thus making unprotected RNA 3' end accessible for binding of the processive TRAMP complex that prevents 3' end processing to resume and allows the exosome to capture the exposed terminus (368, 372). Trf4 may also act on RNAs already polyadenylated by Pap1 (387). Therefore, TRAMP serves also as a scaffold presenting transcripts to the degradation machinery. Interestingly, TRAMP may have an additional role in retaining aberrant transcripts in the nucleus. Air proteins inhibit Hmt1 methyltransferase that modifies mRNA export factors Nab2 and Npl3 (237, 388), and therefore may promote retention of some aberrant transcripts in the nucleus to ensure their degradation by the nuclear surveillance (365). Consistently, polyadenylated RNAs accumulate in the nucleus of cells lacking both Air1 and Air2 (388).

TRAMP in concert with the exosome is also engaged in silencing of some heterochromatic regions in yeast (262, 375, 389). In *S. cerevisiae*, this involves ncRNA transcription and NNS-dependent transcription termination (375, 389). Mechanisms of gene silencing in other organisms, particularly *S. pombe*, also implicate TRAMP-mediated ncRNA degradation, including transcriptional interference, leading to changes in chromatin structure and histone modification (374).

In humans, two MTR4-associated complexes were described, TRAMP and NEXT, that differ is composition, localization and function. The first complex is excluded from nucleoli, and is specifically required for the exosome-mediated degradation of PROMPTs, whereas the second one contains putative homologs of Trf4 (hTRF4-2, also referred to as PAPD5) and Air2 (ZCCHC7) and is restricted to nucleoli, with a postulated function in 3' end adenylation of Pol I transcripts targeted by the exosome (352, 390, 391). An interesting case of a competition between TRAMP and NEXT was recently reported for a biogenesis pathway of human telomerase RNA (hTR) two precursor forms. While the longer variants are preferentially targeted to exosomal degradation by the CBCN complex, the shorter RNA serves as an actual hTR precursor that is oligo-adenylated by TRF4-2 acting in the TRAMP complex, and either processed by PARN (poly(A)-specific ribonuclease) deadenylase or degraded by the exosome (102, 392).

1.4 sn/snoRNP structure and functions

snRNAs comprise a small group of highly abundant ncRNAs that function in pre-mRNA splicing in the nucleoplasm (109). In turn, snoRNAs are mainly localized to the nucleolus and belong to a structurally and functionally diverse families with two major classes being box C/D and box H/ACA (reviewed in (109, 393). Both groups are involved in rRNA processing and modification.

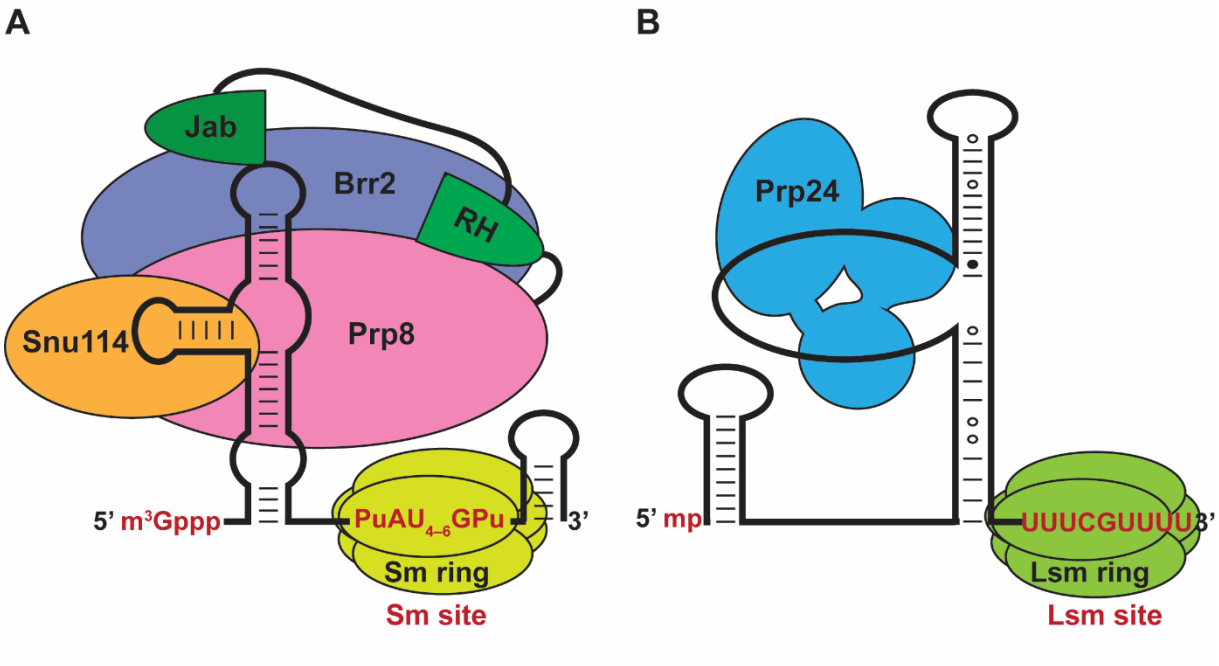

Figure 7. Small nuclear (snRNP) structure and function. *The schematic structure of Sm (**A**) and Lsm class (**B**) snRNPs: (**A**) Structure of yeast U5 snRNA and general location of associated proteins according to* (394, 395) *(**B**) Structure of yeast U6 snRNA and general location of associated proteins according to* (396)*. Conserved sequences are shown in red, dashes represent Watson-Crick base pairing, open and closed circles denote non-Watson-Crick pairing.*

Based on their biogenesis and structure, snRNAs also can be divided into two classes (109). The so called Sm-class, transcribed by Pol II, contains a characteristic 3'stem–loop and is associated with seven Sm proteins, namely B/B', D1, D2, D3, E, F and G (Fig. 7A), which form a heteroheptameric ring structure that binds to the conserved Sm-site-containing $PuAU_{4-6}GPu$ motif (397). This class includes U1, U2, U4 and U5 snRNAs present in all eukaryotes, with additional U7 and minor spliceosome U4atac, U11 and U12 snRNAs in higher eukaryotes (109). In turn, U6 and U6atac (in metazoans), synthesized by Pol III, belong to the Lsm-class, which also contains a 3' stem-loop, but is terminated in a stretch of uridines that bind a distinct heteroheptameric ring of Lsm2-8 (Sm-like) proteins (Fig. 7B) (109). Both groups differ also at their 5' end, as the Sm-class carries a trimethylguanosine (TMG) cap, whereas U6 snRNA has a monomethylphosphate cap (109). Since individual snRNAs

perform different functions in pre-mRNA splicing, they are associated with specific sets of additional proteins (398).

With the exception of mammalian U7 snRNP that functions in histone pre-mRNA 3' processing, U snRNPs are engaged in pre-mRNA splicing forming the core of the spliceosome (109). U1, U2, U4, U5, and U6 snRNAs together with over 150 protein factors make up the major spliceosome. U1 and U2 bind at the 5' splice site and the branch point adenosine, respectively, establishing crucial contacts with pre-mRNA, whereas U5 binds the 3' splice site. Multiple, dynamic RNA- and protein-based interactions contribute to communication between individual snRNPs and pre-mRNA, but intron removal is guided by base-pairing interactions between spliceosomal snRNAs and intron–exon junctions. Recent studies in yeast showed that it is U6 snRNA that catalyses both transesterification reactions by positioning divalent metal ions, which stabilize the leaving groups (399).

SnoRNAs are found throughout both eukaryotes and archaea (snoRNP-like complexes; sRNPs) (393). The eukaryotic H/ACA snoRNAs usually comprise two stems containing H and ACA motifs (Fig. 8A) that are both required for RNP formation. A model based on archaeal sRNPs data predicts that box H/ACA RNP contains two copies of each of the core proteins, i.e. catalytic subunit Cbf5, Nop10, Nhp2 and Gar1 (393). Two Cbf5 proteins probably interact with each other and bind the conserved ACA motifs and the lower stems in snoRNA through their PUA domains (Fig. 8A). Interestingly, mutations that affect a cluster of amino acids in the PUA domain of human Cbf5 (dyskerin) result in dyskeratosis congenital (400). Box H/ACA RNP proteins organize the snoRNA structure and position the target nucleotide in the catalytic centre of Cbf5. Nop10 coordinates snoRNP function, probably interacting with Nhp2, H/ACA RNA and Cbf5, whereas Gar1 seems to be responsible for binding and/or release of the target rRNA (400, 401).

Box C/D snoRNAs contain a C/D motif at the RNA termini and an internal C'/D' motif with the same consensus, though C/D motif seems more conserved (402, 403) (Fig. 8B). The C/D motif is a part of a stem-internal loop-stem structure, known as a K-turn, while the C'/D' often forms a stem-loop structure, named the K-loop (402, 404). The regions between the C and D' or C' and D boxes (guide regions) base-pair with target sites in rRNA. When longer, the extra sequence often folds into a stem structure. Both boxes are associated with one copy of each core proteins, namely RNA-binding protein Snu13, Nop56, Nop58, and the catalytic subunit Nop1 (402). Two parts of the complex are linked by the interaction between Nop56 and Nop58 via their coiled-coil domains (405) (Fig. 8B). Nop56 and Nop58, which

preferentially associate with C/D and C'/D' motifs, respectively, play a direct role in substrate recognition, as they contact guide regions of snoRNAs and rRNA (403, 406).

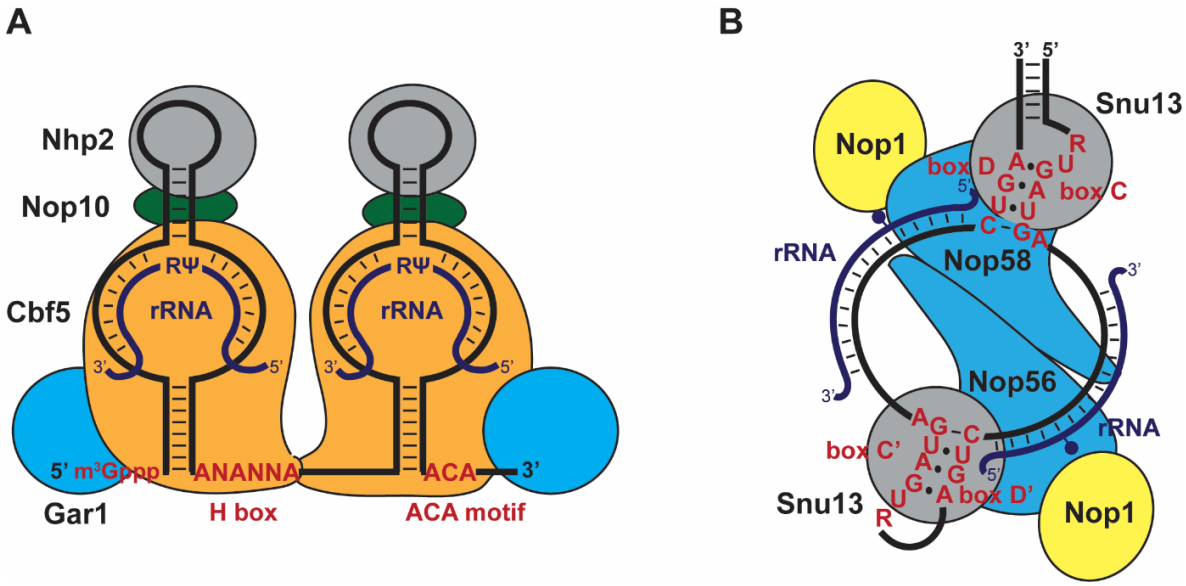

Figure 8. Small nucleolar RNP (snoRNP) structure and function. *Schematic structure of yeast box H/ACA (**A**) and box C/D (**B**) snoRNPs with associated proteins based on archaeal snoRNP-like complexes according to* (393). *The mode of interaction with complementary sequences in rRNA is shown. Conserved sequences are in red, dashes represent Watson-Crick base pairing, and black circles denote non-Watson-Crick pairing. The pseudouridylated or methylated residue is indicated by Ψ (**A**) or a blue circle (**B**), respectively.*

Several additional proteins associate with some snoRNAs in a snoRNP-specific manner (393). The best characterized example is the U3 snoRNP-specific protein Rrp9 (403, 407). It binds the B/C motif in U3 and is essential for U3 recruitment to the processome, a large complex involved in 18S rRNA processing (404, 408–411). Human homologue, hU3-55K, is essential for snoRNP formation and its level regulates the amount of U3 snoRNP during epithelial cell differentiation (412). Several poorly characterized snoRNP-specific proteins associate with yeast snR30 and the Lsm2-7 complex associate with snR5 (393).

Box C/D and H/ACA RNPs catalyse site-specific rRNA modifications, pseudouridylation and methylation, respectively, that take place in the nucleolus (393). Pseudouridines (Ψ) and 2'-O-methyl groups are clustered in functionally important regions of the rRNA, where they stabilize specific RNA structures and are important for ribosome function, which is essential for cell growth (109, 393). snoRNAs are responsible for substrate binding and the choice of the modification site in rRNA. Moreover, they serve as scaffolds that coordinate snoRNP organization. Selection of the target site is achieved through base-pairing between both RNAs. In the case of H/ACA snoRNAs, a pseudouridylation loop in a hairpin structure adjacent to the H or ACA box recognizes the substrate and the unpaired uridine residue is then converted to Ψ by Cbf5 (dyskerin in humans) (413, 414). In turn, the target region in

rRNA interacts with box C/D snoRNA in the vicinity of the D or D' motifs, and the nucleotide located five base pairs away is methylated by Nop1 (fibrillarin in humans) (415–417).

A subset of eukaryotic snoRNPs does not catalyse modification, but is directly engaged in pre-rRNA folding and processing. These include the box C/D snoRNPs U3, that is a part of the rRNA processing complex called the processome, U14, U8, and U22, and the H/ACA snoRNPs snR30/U17 and snR10 (418). Both modifying and processing snoRNPs associate with the target pre-rRNA and have to be subsequently released to enable ribosome biogenesis to proceed and to permit access to other snoRNPs. It has been proposed that ATP-dependent RNA helicases, such as Prp43, are responsible for removing snoRNPs from pre-ribosomes (418).

snoRNPs in higher eukaryotes, in addition to their functions in rRNA biogenesis, modify U6 snRNA (419). Recent studies suggest even broader roles for snoRNPs, including regulation of metabolic stress, alternative splicing, RNA editing and gene expression (109, 393). In humans, telomerase RNA, required for telomere synthesis, belongs to the box H/ACA snoRNA class (420). Some snoRNAs target other RNAs, for example pre-mRNAs, such as brain-specific pre-mRNA encoding serotonin 2C receptor, recognized by snoRNA MBII-52, related to Prader-Willi syndrome (393). MBII-52 is processed to micro (mi)RNA-like fragments that influence alternative splicing and editing of target pre-mRNAs. Many examples of snoRNA-derived miRNAs (sno-miRNAs) were recently discovered in mammalian cells (393). There are also several so-called orphan snoRNAs, which do not appear to target known substrates and could therefore function as miRNA precursors regulating mRNA level. Finally, yeast MRP (mitochondrial RNA processing) ncRNA, component of RNase MRP endonuclease, is also classified as a snoRNA. In yeast, it cleaves rRNA precursor in the nucleus and is also implicated in processing of mitochondrial RNAs to generate primers for mitochondrial DNA replication (421).

1.5 Sn/snoRNA biogenesis: transcription termination, processing and assembly

Eukaryotic sn/snoRNP biogenesis is a highly complex and coordinated process that slightly differs depending on the sn/snoRNA class and gene organization. Whereas most spliceosomal snRNA are encoded by separate genes, gene organization of snoRNAs varies between different organisms: they may arise from independent or polycistronic transcription

units or are excised from pre-mRNA introns (Fig. 9) (422). *S. cerevisiae* snoRNAs are predominately transcribed from independent genes, but some intronic and polycistronic molecules also exist. In turn, more than 90% of human snoRNA genes reside within introns and only few essential snoRNAs involved in pre-rRNA processing are autonomously transcribed (422, 423). Among eukaryotes, both intergenic and intronic snoRNA genes are either standalone or are part of a cluster, although all of eight *S. cerevisiae* intronic snoRNAs are encoded by individual units. Interestingly, most intronic snoRNAs in different known genomes are associated with genes encoding proteins involved in nucleolar function, ribosome structure or protein synthesis. It strongly suggests that biogenesis of partners acting in the same biological process is coordinated (422). There is an evolutionary tendency toward a reduction of the number of independent promoters in snoRNA organization by either clustering of snoRNA coding units (polycistronic transcripts) or colonization of introns (422).

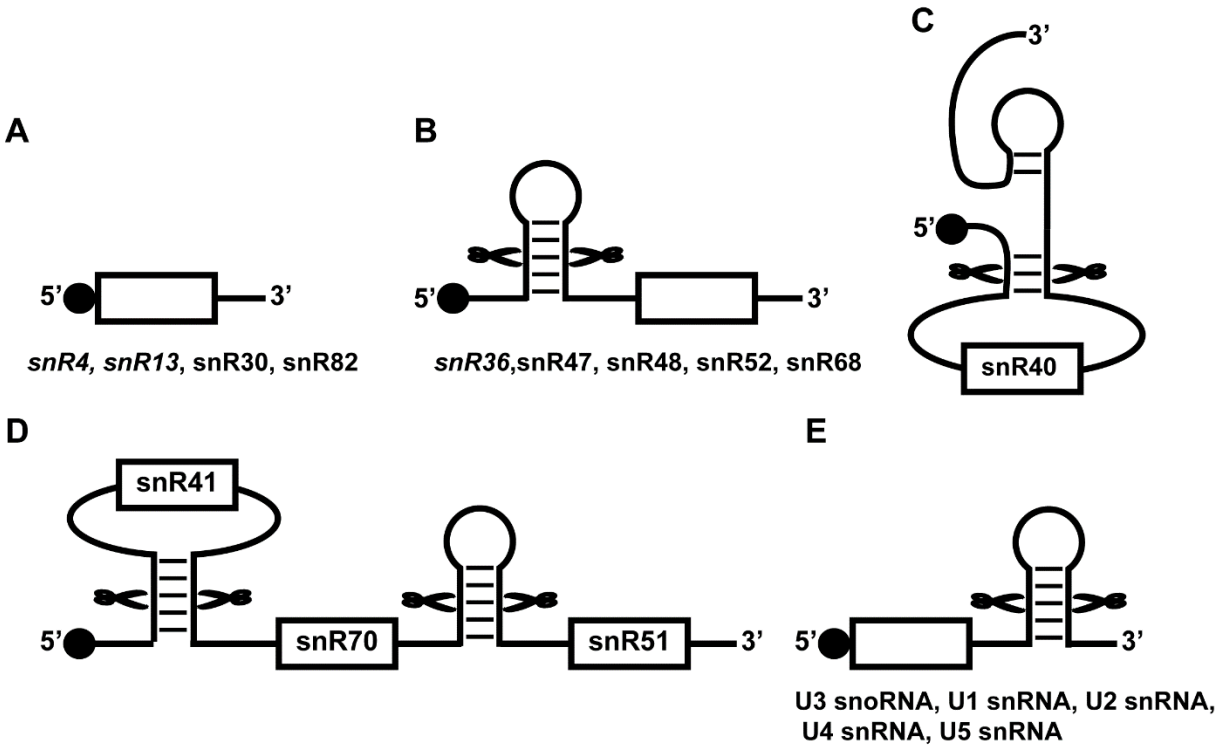

Figure 9. Schematic representation of independently transcribed sn/snoRNA precursor types derived from mono- (A, B, C and E) or polycistronic (D) transcription units according to (249)***.*** *Boxes represent mature sn/snoRNA coding regions, scissors indicate Rnt1 cleavage sites and black circle represents the cap structure. Most of monocistronic box C/D and only a few box H/ACA pre-snoRNAs (e.g. snR36) are cleaved by Rnt1, whereas the majority of box H/ACA and a few box C/D snoRNAs (e.g. snR4 and snR13) retain cap structure at their 5' terminus. Yeast pre-snRNA, except U6, are cleaved by Rnt1 during their 3' end maturation.*

The vast majority of sn/snoRNAs is transcribed by Pol II, except yeast snR52 and U6 snRNA that are Pol III transcripts (158, 422). Single gene encoded sn/snoRNAs are generally transcribed from upstream, poorly characterized promoters containing TATA boxes (424).

Furthermore, promoter regions of sn/snoRNAs in yeast usually comprise A/T-rich elements and binding sites for general transcription factors (Rap1 and Abf1) and, in the case of snoRNAs, also more specific transcription factors (Tbf1) (149, 425). In turn, promoters of U snRNA genes in higher eukaryotes are characterized by a distal sequence element (DSE) that recruits the Oct1 and Sp1 transcription factors and a proximal sequence element (PSE) bound by the specific snRNA activating protein complex (SNAPc) (264).

In the case of snoRNA clusters, individual snoRNAs are released from the polycistronic transcript by Rnt1-mediated endonucleolytic cleavage in the spacer separating each snoRNA (Fig. 9) (131, 149). In turn, processing of intron-encoded snoRNAs is largely splicing-dependent in yeast and mammals (422). However, biogenesis of some snoRNAs in yeast, such as U18, snR39 and snR59, can proceed via two different pathways: major based on conversion of the debranched intron by the lariat-debranching enzyme Dbr1 and minor depending on the endonucleolytic cleavage of pre-mRNA by Rnt1 (131, 150, 158). Most of yeast intronic snoRNAs are flanked by the stem loop recognized by Rnt1, however some of this tetraloops are not canonical (134). The assembly of snoRNPs specific proteins is important for proper release of these snoRNAs from pre-mRNA introns and probably Nop1 acts as one of the chaperones, as its interaction with Rnt1 is necessary for proper cleavage of intron-encoded U18, snR38 and possibly also other box C/D snoRNAs (134).

Transcription termination (see 1.3.2), Rnt1 cleavage of polycistronic or monocistronic pre-snoRNAs or excision from introns generate sn/snoRNA precursors with extended 3’ ends that are further processed by exonucleolytic digestion performed by the exosome/TRAMP complexes (Fig. 10B) (158). Both complexes are co-transcriptionally recruited to pre-sn/snoRNA via their interaction with the NNS complex, at least in the case of independently encoded sn/snoRNAs (see 1.3.2 and 1.3.3). NNS-dependent transcription termination is followed by polyadenylation of released precursor by the canonical poly(A) polymerase Pap1 and the TRAMP component Trf4 (251). The contribution of exosome nuclease activities in pre-sn/snoRNA maturation is not fully established (see 1.3.3.2), but both Dis3 and Rrp6 are engaged in this process (251, 323). Shorter 3’ end extensions are rapidly degraded by Rrp6, whereas the 3’ end of pre-sn/snoRNAs terminated at distant termination elements seems to be processed mainly by the core exosome with the assistance of Rrp6 (251). Rrp6 is more likely responsible for the digestion of final nucleotides, as the progress of the exosome is slowed down by the association of specific protein factors and formation of sn/snoRNPs, that protect the mature 3’ end (see below) (251, 316, 321, 426, 427).

Additional, multiprotein THO complex has been implicated in snoRNA processing in collaboration with the TRAMP and the exosome in fission yeast (428). THO is responsible for coupling between transcription and mRNA processing and negatively regulates the expression of snoRNAs. Both TRAMP and THO complexes are recruited to snoRNA genes, and the latter is required to maintain TRAMP occupancy at snoRNA transcription sites.

snRNAs (except U6 snRNA) and U3 box C/D snoRNA constitute a more specific group in that they are cleaved by Rnt1 at their 3' end prior to exonucleolytic trimming (Fig. 9) and their 3'-extended precursors and cleavage intermediates are stabilized against degradation by binding of Lhp1 (429, 430). However, Rnt1 cleavage is not necessary for accumulation of mature species, as these are also processed from precursors derived from the NNS-dependent terminator (U1, U4 and U5) or PAS (U2) (264). Regardless of the pathway, U1, U2, U4 and U3 are subsequently trimmed by the exosome or alternatively by Rex exonucleases (429, 431). The 3' end processing of two mature forms of U5, U5L and U5S that differ at their 3' ends, is more complex (328). Both functional isoforms can be generated from two distinct Rnt1 intermediate cleavage products, further trimmed by the exosome or Rex2 to U5S and U5S, respectively, but only U5S can be produced in the absence of Rnt1, possibly via the NNS-mediated termination (264). The ratio of both U5 variants is regulated by Rnt1 together with Sen1.

3' end processing of metazoan Pol II transcribed snRNAs proceeds differently and takes place in the nucleus and the cytoplasm (264). The first step involves endonucleolytic cleavage of the nascent transcript that leaves 2–10 nucleotide extension beyond the mature 3' end. This requires the presence of a U snRNA-specific promoter, conserved 3' box, the Integrator complex and Pol II CTD phosphorylated at Ser2 and Ser7 residues (see 1.1.1 and 1.3.2). snRNA precursors are then exported to the cytoplasm using RNA adaptor PHAX that interacts with CBC, and the CRM1/RanGTP complex (see 1.2.2 and 1.2.3). In the cytoplasm, the 3' end is trimmed by unidentified exonuclease (264). In contrast, Pol III-dependent U6 and U6atac snRNAs are processed as typical Pol III transcripts and their termination is driven by the stretch of thymidines at the 3' end (264). Subsequently, the U6-specific uridyl terminal transferase, TUT1 or U6-TUTase, adds a poly(U) tail to the 3' end. This tail is then trimmed to an average of 5 Us by exonuclease Mpn1 that also generates the final 2'-3' cyclic phosphate of the mature U6 snRNA (264).

Processing of 5' termini is more diverse depending on snoRNA (Fig. 10A). In the case of independently transcribed snoRNAs, the 5' end either undergoes endonucleolytic cleavage by

Rnt1 and further exonucleolytic trimming by Rat1 or remains unchanged. In the latter case, their m^7G cap structure is hypermethylated to the TMG cap by methyltransferase Tgs1 (108). In yeast, most box C/D snoRNAs, with the exception of snR4, snR13 and snR45, are cleaved by Rnt1 at their 5' end, whereas this occurs for only three box H/ACA snoRNAs, namely snR36, snR43 and snR46 (Fig. 9). In turn, all Pol II transcribed snRNAs retain their cap structure that becomes hypermethylated. Metazoan snRNA cap trimethylation as well as 3' end processing and snRNP assembly takes place in the cytoplasm, whereas yeast sn/snoRNA and metazoan snoRNA maturation takes place in the nucleus (see 1.2.3). In higher Eukaryotes, snoRNAs transcribed from a separate promoter are transported to the Cajal Bodies thanks to CBC-PHAX/CRM1 interaction (432). After snoRNP assembly, 3' end trimming and cap hypermethylation CRM1 is displaced from the complex by Tgs1 and mature snoRNPs is transport to the nucleolus.

5' end processing of extended pre-snoRNAs requires cap removal as it cannot be hydrolysed by 5'-3' exonucleases. Although all known 5' end extended pre-snoRNA are Rnt1 substrates, *RNT1* deletion does not completely blocks their maturation. The importance of Rnt1 cleavage for this process ranges from being essential (e.g., snR50) to almost dispensable (e.g. Pol III-dependent, uncapped snR52) (120). Therefore, another mechanism enabling exonucleolytic trimming of the 5' end must be employed. It is likely that some decapping enzymes, such as the Dcp1/Dcp2 complex or Nudt proteins, are involved in this pathway that can operate as an alternative or in parallel with Rnt1 cleavage.

5' end of polycistronic snoRNAs that are not first in the cluster is generated by Rnt1 cleavage. In principle, as Rnt1 is implicated in the backup transcription termination mechanism, co-transcriptional Rnt1 cleavage could be followed by degradation of the second and next snoRNAs within the polycistronic unit by Rat1, which could interfere with their processing (141). This may be prevented by the presence of some RNA-binding proteins or transcription factors that protect snoRNA 5' end.

The assembly of snoRNPs is a highly complex and dynamic process that begins co-transcriptionally. Some snoRNP proteins were shown to be recruited to snoRNA genes, influencing their transcription, transcription termination and processing of nascent transcripts. In the case of box C/D snoRNAs, Nop1 associates with snoRNA genes co-transcriptionally and by physical interaction with the APT component Ref2 facilitates its recruitment (316). Similarly, early recruitment of assembly factor Naf1 and core components Cbf5 and Nhp2 to snoRNA genes during transcription is required for efficient assembly of box H/ACA snoRNA

(321, 433). This co-transcriptional association of Naf1 and Cbf5 is facilitated by Pol II CTD Ser2 phosphorylation, as Cbf5 interacts with elongation factor Spt5 and Naf1 directly binds to Pol II CTD (434, 435).

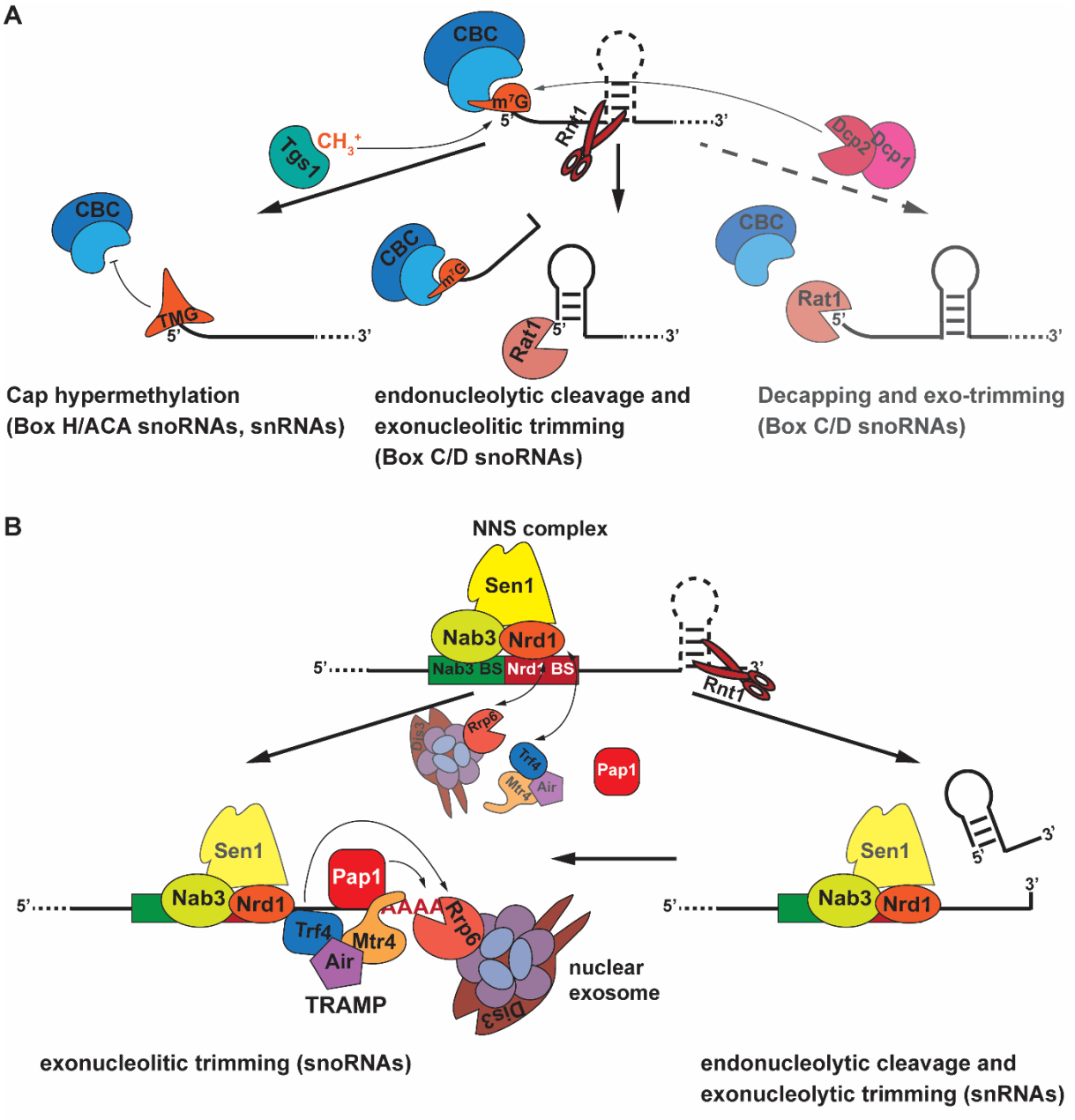

Figure 10. Processing of 5' (A) and 3' (B) ends of sn/snoRNA. *Thin lines with arrowheads indicate catalysed reaction or interaction (arrowheads at both ends), while lines with bar heads indicate inhibition. RNA is drawn as a black line. A stem-loop recognized by Rnt1, present only in some sn/snoRNA precursors, is indicated by dashed lines. Only 5' part (**A**) or 3' part (**B**) of sn/snoRNA precursors is shown. The involvement of Dcp1/Dcp2 decapping complex in the 5' end maturation, proposed by* (120) *as an alternative pathway to Rnt1 cleavage, is indicated by the dashed arrow. Description in the text.*

snoRNP assembly proceeds in several steps with rearrangements that require multiple *trans*-acting auxiliary factors such as Hsp90, R2TP complex and Rsa1 (NUFIP in Metazoa) (436). These chaperones interact with each other and are involved in the assembly of core

proteins of both box C/D and box H/ACA snoRNP. Hsp90 stabilizes Nop17 (also called Pih1), a component of the R2TP complex that also consists of Hsp90 co-chaperone Tah1 and two AAA + helicases Rvb1/Rvb2 (also termed Tip49/Tip48) (436, 437). Rsa1 acts as a scaffold that interacts with the R2TP complex, box C/D specific assembly factor Hit1 and snoRNP core proteins, such as Snu13 and Nop58, which associate as first components with box C/D snoRNAs (438). Binding of Nop56 and Nop58 is assisted by Nop17 that stabilizes Nop56 and increases its affinity to snoRNA, and this allows Nop1 binding to boxes D and D' (439). In turn, Naf1 and Shq1 are involved in box H/ACA snoRNPs assembly. Naf1 facilitates co-transcriptional recruitment of the preformed Naf1-Cpf5-Nop10-Nhp2 complex to the nascent pre-snoRNAs and Shq1 is an RNA mimic that prevents unspecific RNA binding (437). Stability and subcellular localization of Cpf5 (human Dyskerin) is ensured by binding Nhp2 and Shq1. Finally, in the post-transcriptional step, Naf1 is replaced by Gar1 to form the functional box H/ACA snoRNP (437). R2TP complex is required for removal of H/ACA snoRNPs assembly inhibitors from pre-snoRNPs at early stages of the H/ACA snoRNP maturation (440). The assembly of intronic snoRNAs is coupled to pre-mRNA splicing and the position of snoRNA in the intron is critical for this process (441, 442).

The formation of mature spliceosomal snRNPs was extensively studied in higher eukaryotes, where it takes place in the cytoplasm. Sm proteins can spontaneously associate with U snRNA *in vitro*, forming heterooligomeric complexes (D1–D2, B/B0– D3, and E–F–G) that bind to RNA (398). However, *in vivo* U snRNP formation is governed by two complexes, namely the survival motor neuron (SMN) complex and protein arginine methyltransferase 5 (PRMT5) complex (109, 398). SMN, whose reduced expression causes neuromuscular disease called spinal muscular atrophy (SMA), comprises at least eight key subunits (Gemins2–8 and unrip) (443, 444). Sm proteins are sequestered by the PRMT5-complex with important contribution of chloride conductance regulatory protein (pICLn) that directly binds to Sm proteins and PRMT5 (398). Sm proteins B/B0, D1 and D3 are then symmetrically dimethylated, which activates them increasing their affinity to the SMN complex (398). Both complexes form the SMN-/PRMT5 complex and the set of Sm proteins is transferred onto the SMN to be next passed onto snRNA (398). The assembled snRNP is transferred to the nucleus along with the SMN complex and is directed to Cajal Bodies to undergo final maturation, including association of snRNP-specific proteins and snRNA modification (432). SMN, that is a multifunctional complex, may be also involved in box H/ACA snoRNP assembly as it has been shown to interact with Gar1 (437).

Recent studies showed that biogenesis of U4 snRNA also requires Rsa1/NUFIP and the R2TP complex (445). U4 snRNP shares some structural similarity to box C/D snoRNAs and contains Snu13 (15.5K in Metazoa) and Prp31, which similarly to Nop56/58 has a NOP domain. In higher eukaryotes NUFIP directly interacts with the SMN complex, connecting the assembly of Sm and U4 snRNP specific proteins (445).

It is well established that pre-mRNA maturation, including capping, splicing and 3' end formation, occurs co-transcriptionally. The timing of sn/snoRNA processing and its connection with transcription is less obvious, though as described above, efficient and correct 3' end formation of snoRNAs requires co-transcriptional association of snoRNP core factors (316, 321, 426). Components of nascent RNPs may then influence their transcription and processing, but coordination between these aspects for sn/snoRNAs has not been extensively studied. Although the 3' end trimming must occur post-transcriptionally, 5' end cap hypermethylation or Rnt1 cleavage may take place concurrently with RNA synthesis. On the other hand, the interaction between CBC and the NNS complex suggests that CBC remains bound to the nascent sn/snoRNA until the late stages of transcription.

2. RESEARCH OBJECTIVES

The aim of this project was the analysis of the relationship between the processing of 5' and 3' end of sn/snoRNA in yeast *S. cerevisiae*. In particular I wanted to investigate the following aspects:

1. Characteristics of the interaction between CBC and NNS complexes and its contribution to *S. cerevisiae* sn/snoRNA biogenesis;

2. The role of Rnt1 in coupling and synchronizing pre-snoRNA 3' and 5' end processing;

3. The role of the cap methyl transferase Tgs1 and the decapping complex Dcp1/Dcp2 in *S. cerevisiae* sn/snoRNA biogenesis.

3. RESULS

7-methylguanosine 5’ cap is a unique mark of RNA Polymerase II (Pol II) transcripts and its recognition by specific proteins, such as nuclear cap-binding complex (CBC), is essential for regulation of gene expression and RNA stability. CBC is recruited to all Pol II RNAs at the early stages of transcription and is involved in co-transcriptional processing of pre-mRNAs, including transcription termination and 3’ end formation. In contrast, the role of cap structure of ncRNAs is not fully understood.

Yeast sn/snoRNAs never leave the nucleus and the importance of cap and CBC for their biogenesis has not been conclusively demonstrated, despite strong indications that CBC interacts with the ncRNA-specific termination NNS complex (103). Preliminary data revealed that when 5’ processing is inefficient snoRNA precursors do not possess mature 3’ ends (Grzechnik, unpublished results). This 3’ processing defect is further increased when CBC is missing suggesting its involvement in a quality control mechanism that coordinates 5’ and 3’ end processing. The aim of this project was to thoroughly analyse the role of CBC and endonuclease Rnt1 in coordinated processing of sn/snoRNA 5’ and 3’ termini. Since maturation of sn/snoRNAs shares several common features I concentrated mainly on selected single gene-encoded box C/D snoRNAs.

3.1 Specificity and efficiency of new cap analogue modified affinity resins

Affinity chromatography is a useful method for protein purification and it can be successfully applied for protein interaction studies. Working on my master degrees in chemistry and biology I synthesized three novel affinity resins with attached cap analogues (Fig. 11A): m^7GpCH_2pp-Sepharose (**2**), m^7GpppA-Sepharose (**3**) and m^7GpCH_2ppA-Sepharose (**4**) (446). Two of these resins (**2** and **4**) contain methylene group replacing the pyrophosphate oxygen atom closest to the m^7G moiety (between β and γ phosphate groups). This modification was introduced to specifically inhibit the decapping activity of DcpS, which hydrolyse pyrophosphate bond in this position. During the course of the PhD project I examined the properties of these resins and exploited them to develop the binding assay to study the interaction between CBC and NNS complexes. Results presented in this section were published in the RNA journal (446).

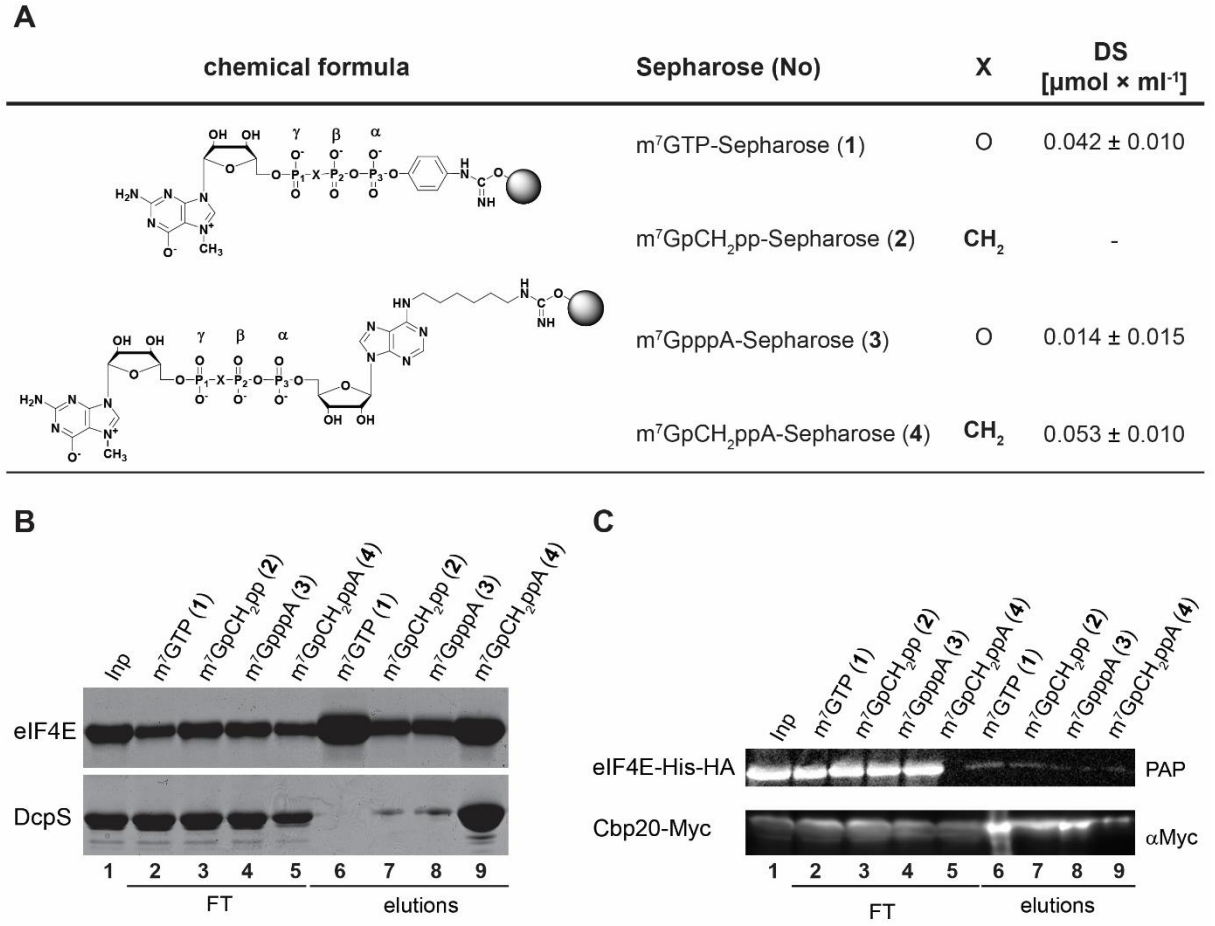

Figure 11. Properties of new cap analogue modified affinity resins: *(**A**) Structures of the m^{7}GTP-Sepharose* (447) *and affinity resins used in this study* (446)*. DS values (degree of substitution) were determined by enzymatic digestion. (**B**) Binding of recombinant mouse eIF4E$^{(28–217)}$ and human DcpS to resins (**1-4**). Coomasie blue stained SDS-PAGE of input (inp, lane 1), flow-through (FT, lanes 2-5), and bound fractions eluted with 1 mM m^{7}GTP (lanes 6-9). (**C**) Western blot of eIF4E-His$_6$/HA/ProtA and Cbp20-Myc purified from yeast extract using resins (**1-4**), eluted with 1 mM m^{7}GTP and detected with PAP and anti-Myc antibodies, respectively. Western blot was performed on total lysate (Inp, lane 1), flow through (FT, lanes 2-5) and elutions from each resin (lanes 6-9).*

The degrees of substitution (DS) to the resin were determined as described previously (447). Ligands were released from matrices using two enzymes: alkaline phosphatase and phosphodiesterase I, and their concentration was measured spectrophotometrically at 260 nm. The degree of substitution for Sepharose 2 could not be calculated using this method, as this analogue was not cleaved by either of the enzymes. The DS values for all tested resins are presented in Fig. 11A. They were within the same range but lower than those described before (447).

The efficiency of cap binding protein purification was evaluated using recombinant mouse eIF4E$^{(28-217)}$ and human DcpS. Proteins were incubated with equal volumes of m^{7}GpCH$_2$pp-Sepharose (**4**), m^{7}GpppA-Sepharose (**3**), m^{7}GpCH$_2$ppA-Sepharose (**2**) and control m^{7}GTP-Sepharose (**1**), synthetized as described in (447). Bound fractions were eluted with free

m^7GTP cap analogue (1mM) and precipitated from pooled eluates using pyrogallol red (448). Binding efficiency of m^7GpCH_2ppA-Sepharose (**4**) were comparable to the standard m^7GTP-Sepharose (**1**) for mouse eIF4E$^{(28-217)}$ (Fig. 11B), while two other resins (**2**, **3**) showed significantly lower binding affinities (Fig. 11B). As expected, only non-hydrolysable, methylene modified resins (**2** and **4**) bound human DcpS. However, the only resin that bound this pyrophosphatase with high efficiency was the m^7GpCH_2ppA-Sepharose (**4**) with a hexylene spacer linking the cap to the resin (Fig. 11B).

The applicability of new resins for purification of cap binding proteins from a more complex biological material, such as cellular extracts, was studied before (446). To assess the relative efficiency of protein binding to different cap analogues more quantitatively, I performed similar experiments using yeast strains expressing tagged yeast cap-binding proteins, eIF4E and Cbp20. eIF4E was expressed from a plasmid under the control of the pGAL promoter as a C-terminal fusion with a triple affinity tag consisting of a His_6-HA epitope, a protease 3C cleavage site and the IgG domain (ZZ) of protein A (450). Cbp20 was expressed under its endogenous promoter in fusion with the C-terminal Myc epitope (86). The level of eIF4E was comparable in all four eluates (Fig. 11C), which is consistent with the MS-MS data (446), but differed significantly from binding studies with mouse eIF4E$^{(28-217)}$ (Fig. 11B). This discrepancy probably results from structural differences between yeast and mammalian eIF4E that can affect protein-cap complex stability (451, 452). In turn, Cbp20 bound with a comparable high affinity to the control m^7GTP-Sepharose (**1**) and the newly synthesised m^7GpppA-Sepharose (**3**). The level of Cbp20 in eluates from m^7GpCH_2pp-Sepharose (**2**) and m^7GpCH_2ppA-Sepharose (**4**) was much lower, in agreement with the MS-MS data (446). Resins **1** and **3** were therefore chosen for binding assays to evaluate the interaction between CBC and NNS complexes.

3.2 CBC directly interacts with the Nrd1/Nab3 complex

CBC proteins copurify with components of the NNS complex (103), but the character of this interaction has not been established. To confirm direct binding between these proteins I performed pull-down assay using yeast cellular extracts expressing Cbp80-Myc and Cbp20-Myc and the recombinant Nrd1/Nab3 complex. To this end I purified Nrd1$^{(1-548)}$-His/Nab3$^{(191-556)}$-MBP, lacking large N- and C-terminal parts of Nab3, but containing all conserved motifs with single RRMs in both Nrd1 and Nab3, as well as the CTD interaction domain (CID) of Nrd1 (274). This experiment showed that CBC copurified with the Nrd1/Nab3 complex on Ni-NTA-Agarose (Fig. 12A and B), whereas CBC alone was not retained on the beads (Fig.

12A). In contrast, Rad30-Myc, a nuclear DNA polymerase eta, used as a negative control, did not bind Ni-NTA-Agarose either in the presence or absence of Nrd1/Nab3 proteins (Fig. 12C). Since both complexes contain RNA binding proteins, their copurification from the extract can simply result from simultaneous interaction with RNA molecules. To exclude this possibility, pull-down was repeated using RNase A-treated extract (experiment performed under my supervision by the master student Dorota Adamska). As expected, RNase treatment did not affect CBC-Nrd1/Nab3 interaction (Fig. 12B, compare lanes 4 and 6), strongly suggesting it is not mediated through RNA.

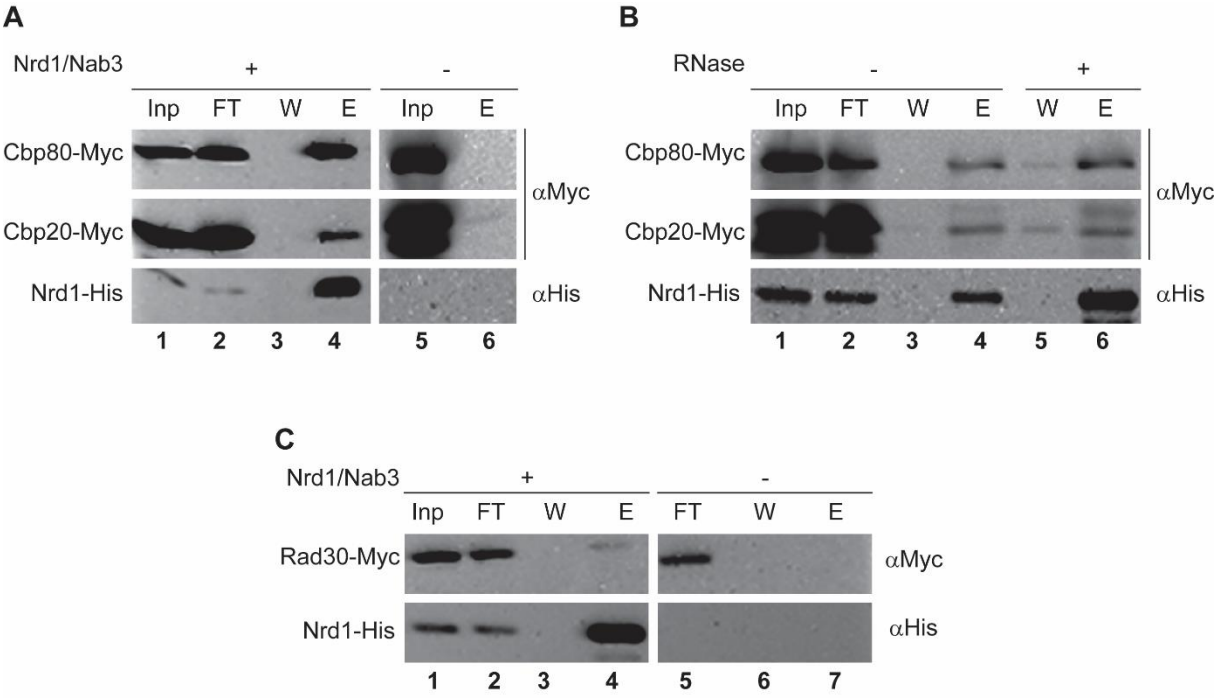

Figure 12. CBC interacts with the Nrd1/Nab3 complex. *(**A**) Pull-down assay: the mixture of Cbp20-Myc and Cbp80-Myc extracts incubated with or without Nrd1$^{(1-548)}$-His/Nab3$^{(191-556)}$-MBP and purified on Ni-NTA-Agarose. Western blot analysis: input (Inp, lanes 1 and 5), flow through (FT, lane 2), last wash (W, lane 3) and elution (E, lanes 4 and 5). (**B**) Pull-down assay using RNase A treated extracts, description as for **A**. Prior to incubation with the Nrd1/Nab3 complex, the extract was treated with RNaseA (eluate in lane 6). (**C**) Control pull-down assay: Rad30-Myc extract incubated with or without Nrd1$^{(1-548)}$-His/Nab3$^{(191-556)}$-MBP and purified on Ni-NTA-Agarose, description as for **A**.*

Other putative Nrd1-interacting proteins identified in the Nrd1-TAP isolation included Pol II subunits and components of numerous RNA processing complexes (103). To determine whether CBC-NNS interaction is direct or mediated by other factors I performed *in vitro* binding using purified complexes. Because I could not obtain stable CBC expression in bacterial system, CBC was purified from yeast strains overexpressing Cbp20-His$_6$-HA-3C-ZZ or Cbp80-FLAG-Strep-tagII$_2$ under the control of the P_{GAL} promoter. The isolated complex was contaminated with low amounts of protease 3C (Fig. 13), but no subunits of NNS, Pol II

or exosome complexes were detected as determined by mass spectrometry (data not shown). Cbp20-His_6-HA-ZZ/Cbp80-FLAG-Strep-tagII$_2$ and Nrd1$^{(1-548)}$-His/Nab3$^{(191-556)}$-MBP complexes were incubated and purified using the m^{7}GTP-Sepharose (**1**) and m^{7}GpCH$_2$ppA-Sepharose (**3**) (446). Both complexes were present in eluates, although the Nrd1/Nab3 complex was less abundant (Fig. 13). At the same time proteinase 3C and other contaminants were strongly depleted. Upon incubation of Nrd1$^{(1-548)}$-His/Nab3$^{(191-556)}$-MBP with both resins, Nrd1 itself showed only a weak binding (Fig. 13 lane 9 and 10), demonstrating that Nrd1/Nab3 copurifies with CBC through their interaction. The results obtained using both resins were comparable, although the standard m^{7}GTP-Sepharose (**1**) was more efficient and specific. These results demonstrate that the CBC-Nrd1/Nab3 interaction is direct, not dependent on other proteins or RNA. Importantly, binding of the cap by CBC does not interfere with this process, which indicates that CBC-Nrd1/Nab3 may engage the cap-bound CBC *in vivo*.

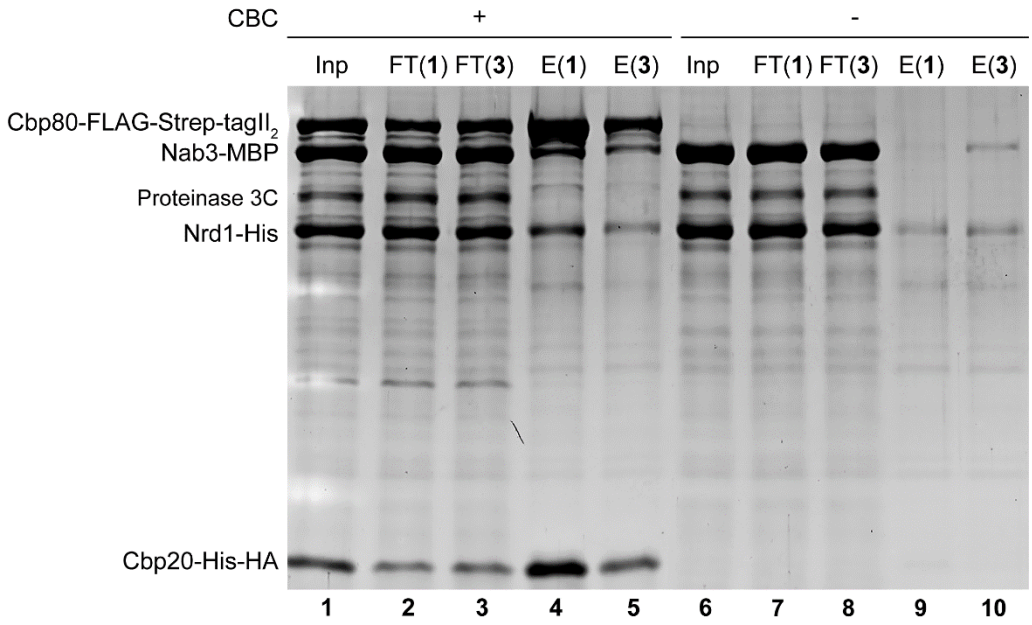

Figure 13. CBC interaction with the Nrd1/Nab3 complex is direct. *Interaction in vitro using purified complexes: Cbp80-FLAG-Strep-tagII2/Cbp20-His-HA and Nrd1$^{(1-548)}$-His/Nab3$^{(191-556)}$-MBP. CBC complex, purified from yeast, was contaminated by Proteinase 3C, Pol II subunits or RNA, but not with the NNS complex, as determined by MS analysis. Interacting proteins were purified using m^{7}GTP-(**1**) and m^{7}GpppA-(**3**) Sepharoses. Control binding of the Nrd1/Nab3 complex (without CBC) is presented on the right. SDS-PAGE analysis: input (inp, lanes 1 and 6), flow through (FT, lanes 2-3 and 7-8) and elution (E, lanes 4-5 and 9-10).*

3.3 CBC is recruited to Pol II transcribed sn/snoRNA genes

Recruitment of CBC at the early step of transcription was reported for protein coding genes but not for ncRNA genes (33, 86, 88). To determine CBC association with much shorter sn/snoRNA genes I optimized chromatin immunoprecipitation (ChIP) using extensive

sonication of formaldehyde-crosslinked chromatin to obtain fragments ~200 bp in length (10). The occupancy of Cpb20, Cbp80 and Pol II was analysed in Cbp20-Myc and Cpb80-Myc strains for chosen genes, using anti-Myc and anti-Rpb3 antibodies. Selected sn/snoRNAs represented different classes and gene organization and were relatively well separated from other transcription units or expressed at a much higher level than neighbouring genes, according to available high-throughput data (8, 10). This group included independently encoded box C/D snoRNAs snR47, snR48 and snR68, whose precursors are cleaved by Rnt1 at their 5' ends, and snR45 not processed by Rnt1; polycistronic box C/D snoRNA unit, snR41-snR70-snR51, where individual snoRNAs are separated by Rnt1 cleavage sites; box H/ACA snoRNAs snR30 and snR82 not cleaved by Rnt1; and U1 and U5 snRNAs processed by Rnt1 at their 3' ends. Pol III transcribed box C/D snoRNA snR52 was used as a negative control. CBC and Pol II (Rpb3) profiles for these genes are shown in Figures 14 - 16.

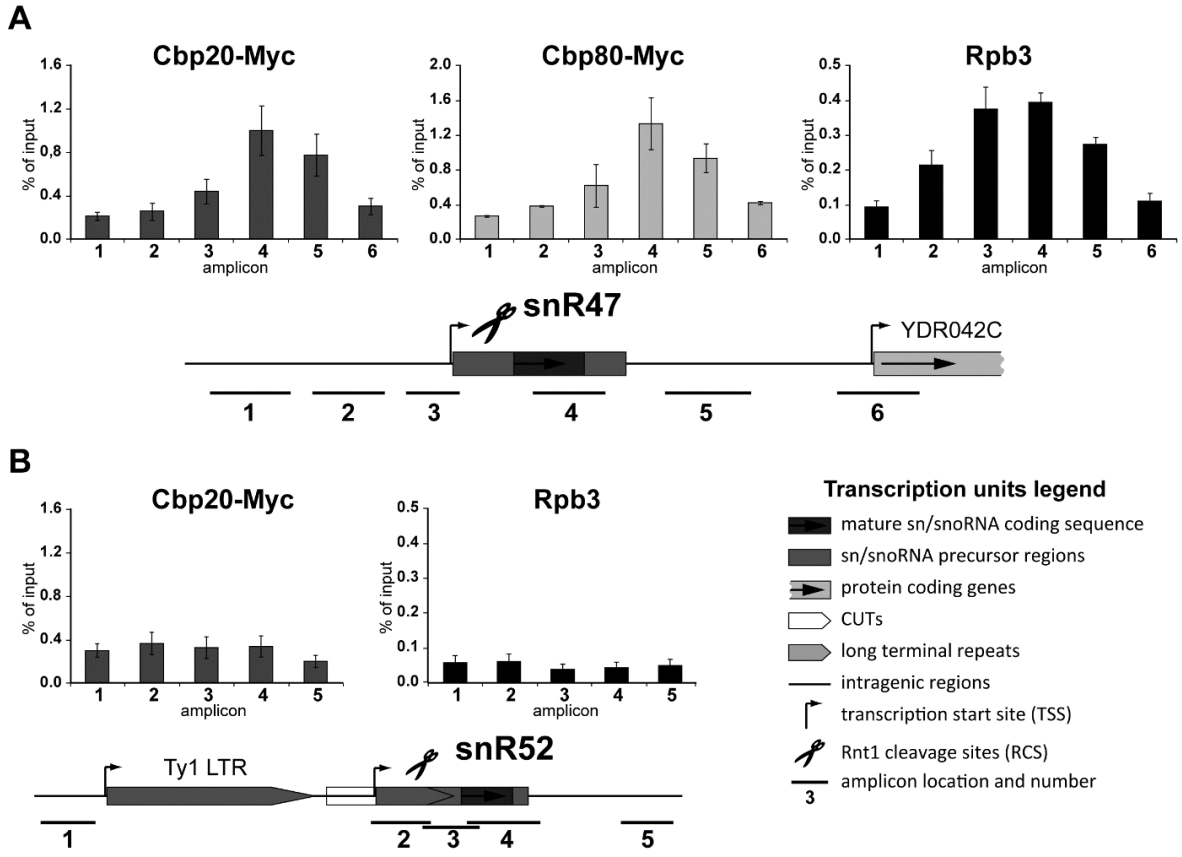

Figure 14. CBC is co-transcriptionally recruited to Pol II transcribed sn/snoRNA genes. *Chromatin immunoprecipitation (ChIP) for CBC (Cbp20-Myc or Cbp80-Myc) and Pol II (Rpb3). CBC and Rpb3 profiles are shown along box C/D snR47 (**A**) and Pol III transcribed snR52 (**B**) genes. qPCR analysis of the ChIP output in CBP20-Myc (dark bars) or CBP80-Myc (light bars) strain using an anti-Myc (for Cbp20 or Cbp80 proteins) and anti-Rpb3 (black bars) antibodies, respectively. Data is presented as the percentage of input. The scale on the diagrams is adjusted to the maximal Cbp20 or Rpb3 occupancy at snR47. Error bars indicate standard deviation from at least three independent experiments. Schemes shown below the diagrams represent distribution of amplicons (marked 1-6), sn/snoRNA gene organization and adjacent transcription units. All symbols used here and in subsequent figures are explained in the Transcription unit legend on the bottom-right.*

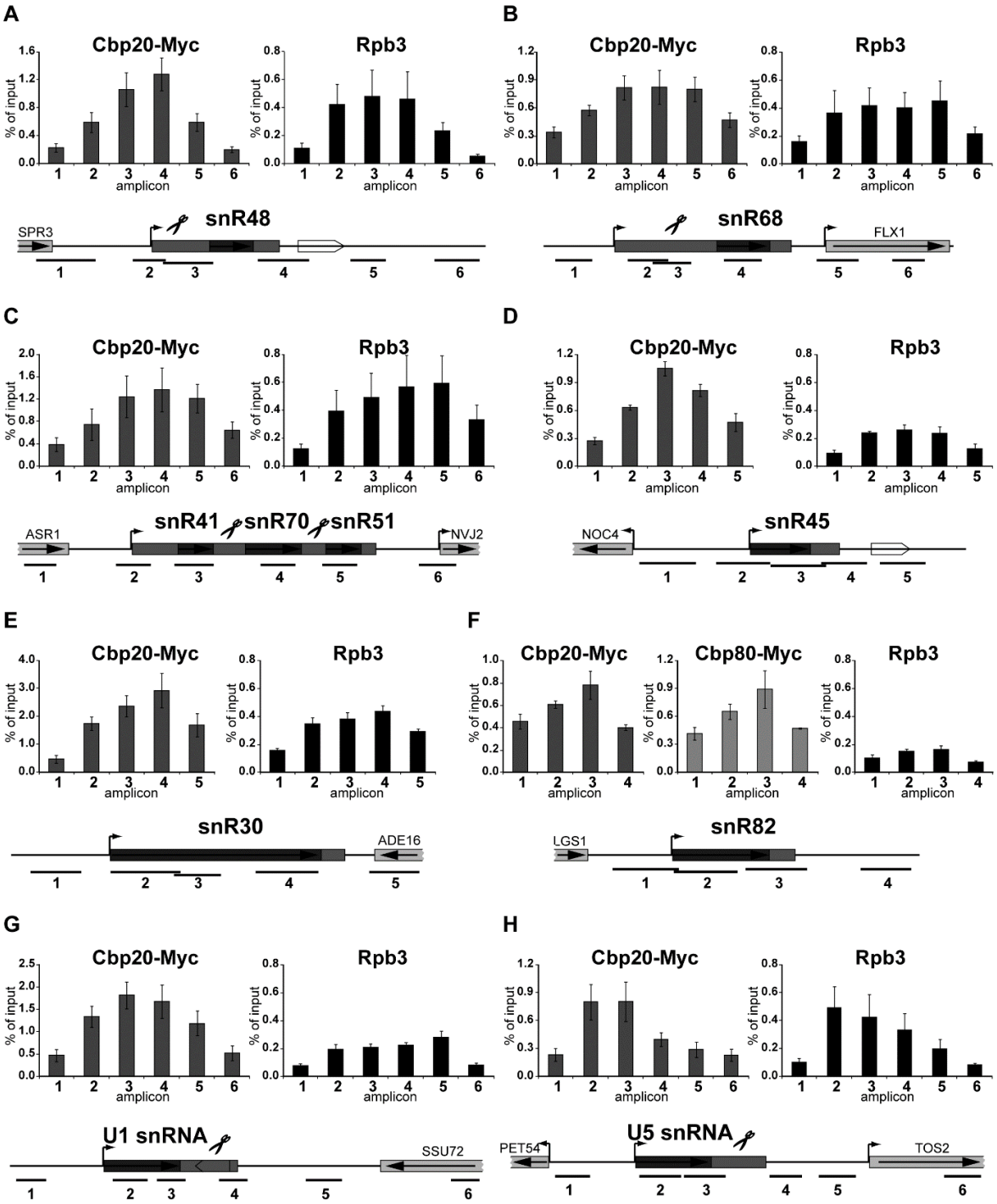

Figure 15. CBC is co-transcriptionally recruited to Pol II transcribed sn/snoRNA genes. *CBC (Cbp20-Myc or Cbp80-Myc) and Pol II (Rpb3) profiles at box C/D snoRNA genes (**A**-**D**), box H/ACA snoRNA genes (**E** and **F**) and snRNA genes (**G** and **H**). Description as for Figure 14. The scale on the diagrams is adjusted to the maximal Cbp20 or Cbp80 occupancy at the given gene and set for 0.8% for all anti-Rpb3 IPs.*

As expected, CBC is recruited to all tested Pol II transcribed sn/snoRNAs, but not to snR52, for which the signal is similar to that observed for the noncoding control region (Fig. 14-15). The slight enrichment of CBC and Pol II signals for snR52 for amplicons 2-4 may result from the expression of CUT that overlaps with pre-snR52 5' end (Fig. 14B). Cbp20 and

Cbp80 profiles are similar (Fig. 14A and 15F), therefore for most genes only Cbp20 occupancy was assigned. As in the case of protein coding genes, CBC recruitment occurs early in the transcription cycle reaching the full occupancy about 200 bp downstream of the transcription start site (TSS) (33). CBC profiles at the 3' end of sn/snoRNA genes follow Pol II profiles with the prominent decrease downstream of the transcription termination region. The discrepancy between Cbp20 and Rpb3 profiles at the 3' end of U1 snRNA may result from Pol II pausing at the terminator region of the *SSU72* gene (Fig. 15G). Similarly, slightly elevated Rpb3 level at the 3' end is observed for snR30 situated in the close vicinity of *ADE16* (Fig. 15E). These results suggest that CBC remains associated with sn/snoRNA genes until termination takes place, regardless of the presence of Rnt1 cleavage site in the precursor. Noteworthy, CBC level at analysed genes is similar for all sn/snoRNA classes, regardless of the presence or absence of 5' Rnt1 cleavage site (Fig 14-15). It is also not correlated with the Rpb3 level, but it has a slight tendency to reach higher values at longer transcription units (Fig 15C and E).

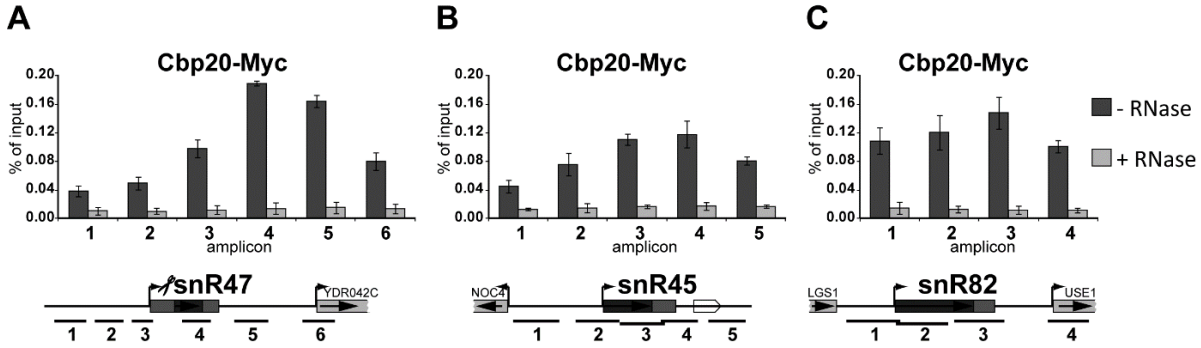

Figure 16. CBC recruitment is dependent on RNA. *Cbp20 profiles at snR47 (**A**), snR45 (**B**) and snR82 (**C**). The results of qPCR analysis of the ChIP output with (light bars) or without (dark bars) RNase treatment. Description as for Figure 14. The scale on the diagrams is set for 0.2%.*

To test if CBC presence at sn/snoRNA genes depends on RNA I performed ChIP using RNase-treated extracts for selected snoRNAs (snR47, snR45 and snR82) (Fig. 16A, B and C). ChIP signals for RNase-treated samples were near background levels, although at least two orders of magnitude higher than those obtained for the wild-type extract (BY4741 strain, data not shown). This clearly indicates that CBC is bound to nascent sn/snoRNA precursors. In general, observed occupancy is lower than in previous ChIP experiments (Fig. 14-15) due to differences in the protocol, especially shorter incubation time with the antibody.

3.4 Rnt1 is co-transcriptionally recruited to sn/snoRNA genes

Association of Rnt1 with rDNA was shown to play an important role in co-transcriptional pre-rRNA processing (132). It has been suggested that endonucleolytic cleavage of snoRNA precursors by Rnt1 occurs co-transcriptionally, based on the enrichment of Rnt1 near the snR190-U14 gene (132). It is more likely in the case of processing in the 5' precursor region that has the potential to fold early on into the hairpin structure recognized by Rnt1. On the other hand, association of CBC with sn/snoRNA precursors until the termination region indicates that the cleavage may be delayed and take place shortly before or even following termination. To determine whether Rnt1 is co-transcriptionally recruited to sn/snoRNA genes and when this event actually occurs I performed ChIP analysis using Rnt1-Myc strain for Rnt1-cleaved sn/snoRNAs: snR47, snR48, snR68, snR41-snR70-snR51 and U1, U5 snRNAs, as well as for snR45, snR30 and snR82 not processed by Rnt1 and Pol III transcribed snR52.

For all Pol II transcribed and Rnt1-processed sn/snoRNAs, except U5, I observed significant enrichment of Rnt1 at the 3' ends of genes and their termination regions (Fig. 17). Notably, Rnt1 remains associated with snoRNA genes until very late stages of transcription, with the highest occupancy observed in regions where Pol II level already drops, which is especially visible in the case of snR48 (Fig. 17B, compare with Fig. 15A) and snR41-snR70-snR51 (Fig. 17D, compare with Fig. 15C). Although Rnt1-Myc signal is weaker compared to that of Cbp20-Myc, and there is only approximately 2 fold enrichment relative to adjacent noncoding regions of each gene, Rnt1 profiles for all tested Pol II transcribed snoRNAs are consistent. In contrast, Rnt1 profile for the Pol III transcribed box C/D snoRNA snR52 is significantly different (Fig. 17J), with the highest occupancy at the 5' region of the precursor, where Rnt1 cleavage site is present (120), and dropping towards the 3' end of the gene. Association of Rnt1 with snRNA genes was also distinct, as expected, since their precursors are cleaved at their 3' termini that become accessible for recognition at the very end of their transcription. There is no detectable enrichment of Rnt1-Myc at U5 (Fig. 17F) and a quite discreet Rnt1-Myc peak downstream of U1, overlapping with the Pol II profile 3' of the terminator region (Fig. 17E, compare with Fig. 15G). In addition, no Rnt1 enrichment was detected at box C/D snR45 and box H/ACA snR82 that are not processed endonucleolytically (Fig 17G and I). However, slight Rnt1 signal was observed at the 3' end of the box H/ACA snR30, not reported to be cleaved by this enzyme (Fig. 17H). Rnt1 level is similar for most tested box C/D snoRNA genes, with the exception of snR68 (Fig. 17C) that showed three fold higher maximal occupancy. This gene was therefore chosen for subsequent ChIP, northern

blot and cRT-PCR analyses. On the other hand, Rnt1 level measured along snR48 gene was visibly lower (Fig. 17B), possibly due to a weaker interaction of Rnt1 with the noncanonical tetraloop (see 1.2.4.1).

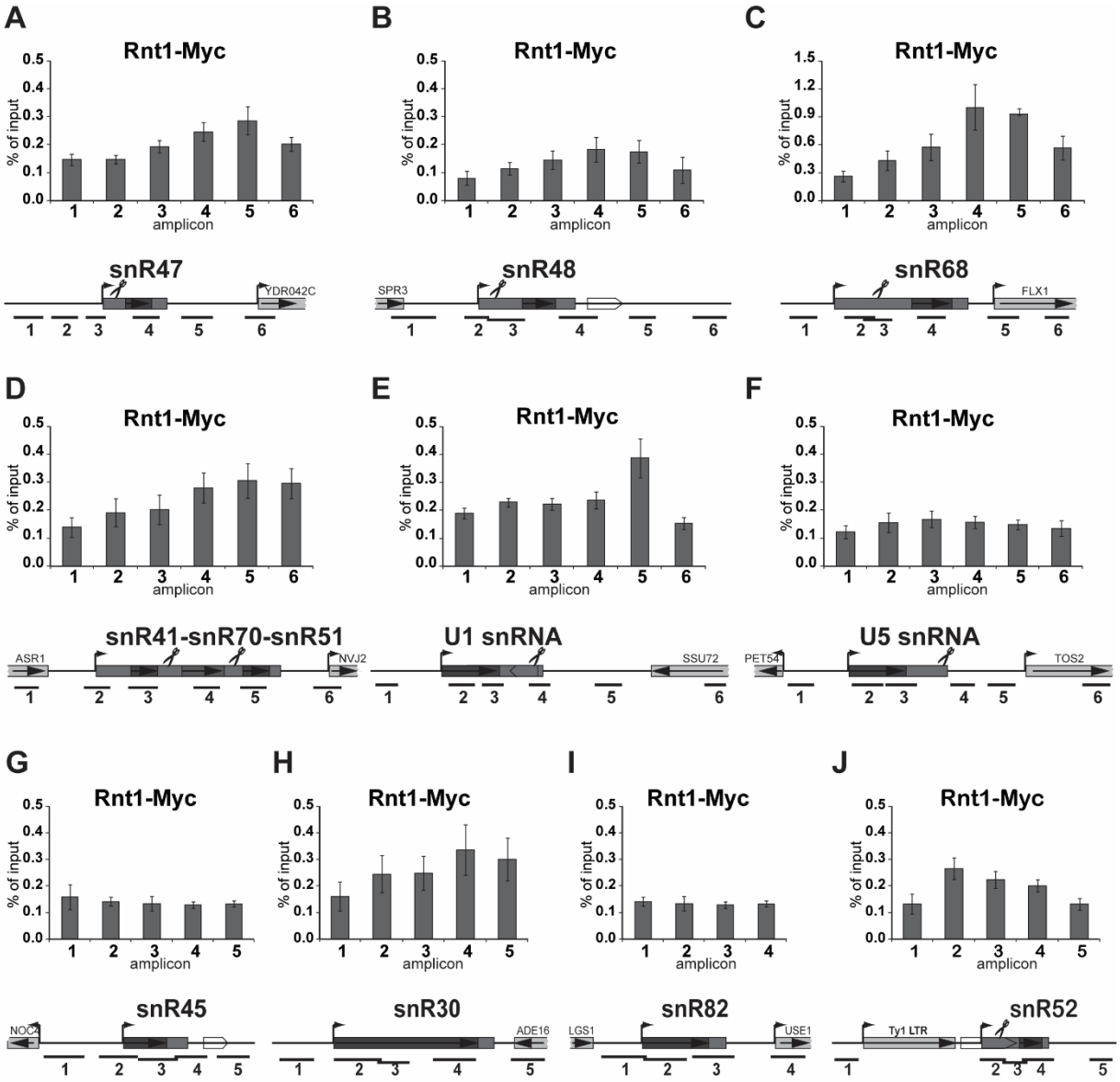

Figure 17. Rnt1 is co-transcriptionally recruited to sn/snoRNA genes. *(**A-J**) Rnt1 profiles at different sn/snoRNA genes. qPCR analysis of the ChIP output in RNT1-Myc strain using anti-Myc antibodies. Description as for Figure 14. The scale on most diagrams is set for 0.5%, with the exception of snR68 (**C**), which shows the highest level of Rnt1 occupancy.*

Since ChIP results for Rnt1-Myc performed after RNase treatment were inconclusive due to the low signal and high background, to determine whether Rnt1 recruitment to sn/snoRNA genes depends on RNA I carried out ChIP of Rnt1-Myc in the *snR68Δ* strain expressing either wild-type or mutated copy of plasmid-borne snR68. The mutated version carries two substitutions in the tetraloop recognized by Rnt1 (AGGA→ACAA), which disable Rnt1

binding and endonucleolytic cleavage (124). The results of this ChIP experiment showed that mutations within the snR68 tetraloop (snR68mut) prevent Rnt1 recruitment to its gene, while the occupancy at other snoRNA genes was not affected (Fig. 18A and B). This also confirmed that co-transcriptional recruitment of Rnt1 involves the recognition of the tetraloop. Rnt1 profile for the episomal copy of snR68 differs from that of genomic snR68 in the *RNT1-Myc* strain (Fig. 17C), probably due its transcription from the plasmid.

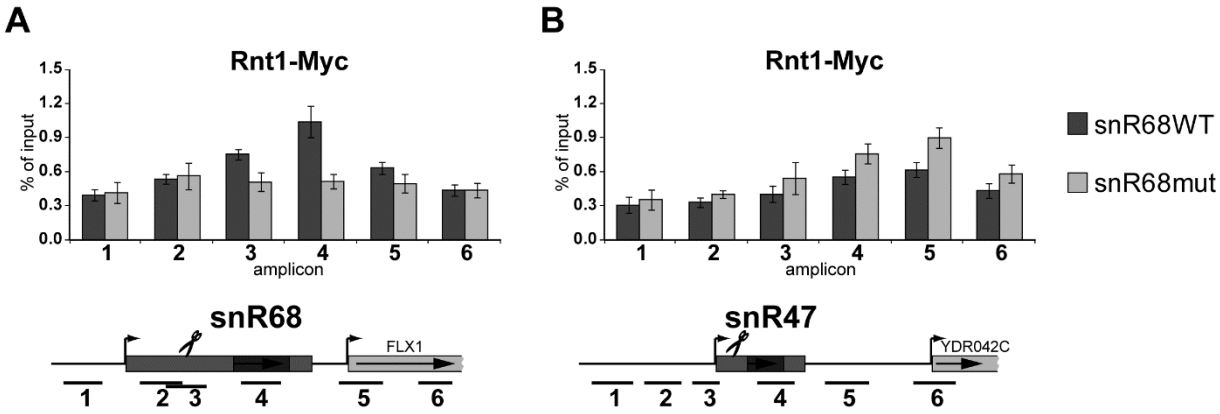

Figure 18. Rnt1 recrutment depends on tetraloop recognition. *Rnt1 profiles at snR68 (**A**) and snR47 (**B**). qPCR analysis of the ChIP output in RNT1-Myc snr68Δ pRS415-snR68WT (dark bars) and RNT1-Myc snr68Δ pRS415-snR68mut strains using an anti-Myc antibody. The mutated version of snR68 carries two substitutions in the tetraloop recognized by Rnt1 (AGGA→ACAA). Description as for figure 14. The scale on the diagrams is set for 1.5%*

Rnt1 was shown to interact directly with Sen1 and to associate with the NNS complex, probably via Sen1 (36, 103). To assess whether Rnt1 recruitment to snoRNA genes requires interactions with other protein factors, I performed ChIP for Rnt1-Myc in the *sen1-K128E* mutant, where binding between Sen1 and Rnt1 is disrupted (37, 292). As a control Rnt1-Myc *sen1-R302W* strain was used with impaired Sen1-Pol II CTD interaction, but retained Sen1-Rnt1 connection. Rnt1 profiles in both strains showed no significant differences and were comparable to that for the wild-type (Fig. 19). This result indicates that Sen1 has no significant contribution to Rnt1 association with snoRNA genes.

3.5 Lack of CBC delays Nrd1 recruitment to sn/snoRNA genes

The NNS complex is recruited to sn/snoRNA genes through its interaction with Pol II CTD and direct binding of Nrd1 and Nab3 to RNA (272, 273, 275). It has been shown that Nrd1 interacts with the CTD phosphorylated at Ser5 (Ser5-P), typical for Pol II at the 5' end of genes (272), whereas Sen1 specifically binds Ser2-phosphorylated CTD (Ser2-P), which level increases towards the 3' end of genes (37).

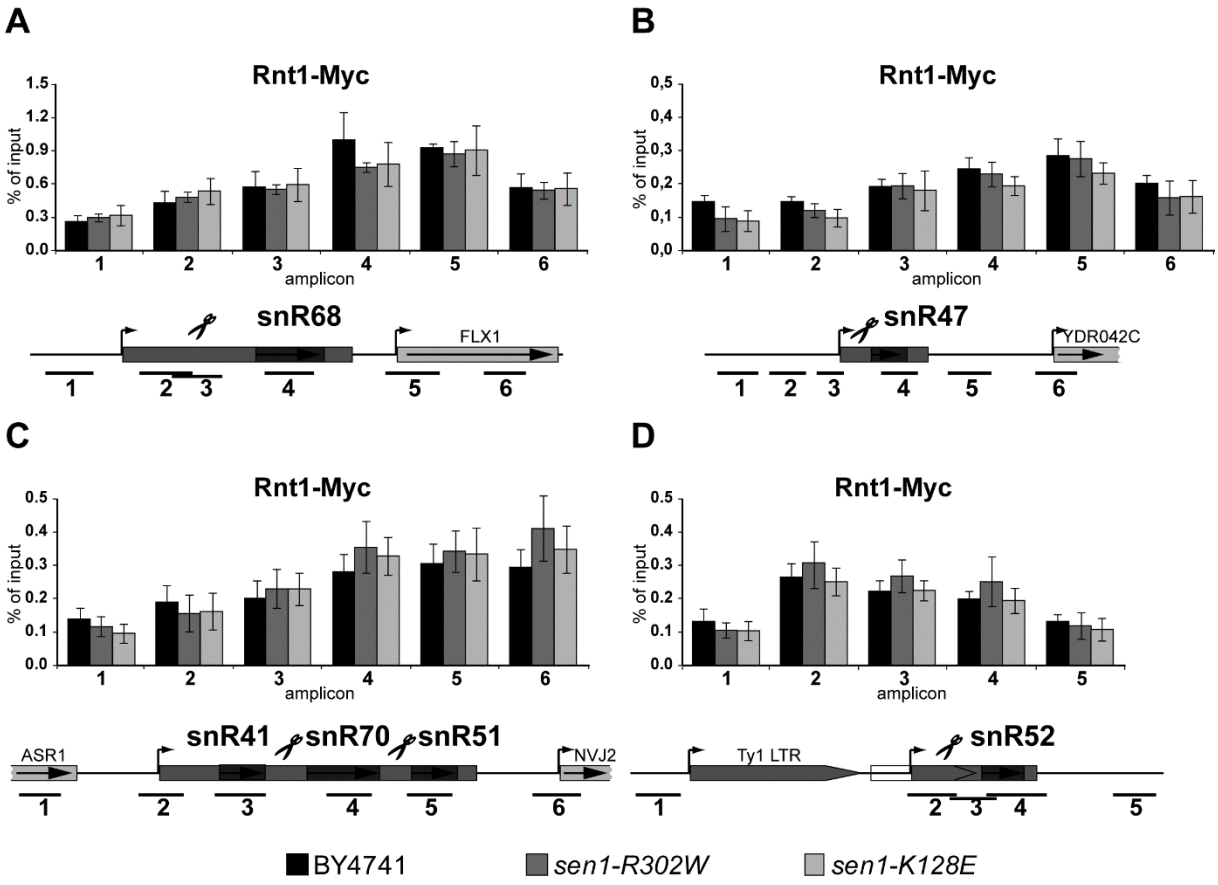

Figure 19. Rnt1 recruitement to sn/snoRNA genes does not depend on interaction with Sen1. *(**A**-**D**) Rnt1 profiles at chosen snoRNA genes. qPCR analysis of the ChIP output in the wildtype (RNT1-Myc in BY4741), RNT1-Myc sen1-R302W and RNT1-Myc sen1-K128E strains using anti-Myc antibodies. Description as for Figure 17.*

To test whether association of CBC and Nrd1 mutually affects one another I performed ChIP analysis for different classes of sn/snoRNAs for Cbp20-Myc in the *GAL::Nrd1-HA* strain before and after Nrd1 depletion, and for Nrd1-HA in the *cbp80Δ** mutant. Nrd1 depletion had no direct effects on CBC association with tested snoRNA genes (data not shown). In contrast, in the absence of CBC, Nrd1 occupancy at sn/snoRNA genes was slightly but significantly lower than in the wild-type strain (Fig. 20). The strongest effect was observed for snR68, snR45 and U1 (Fig. 20A, C and E), and it was not caused by the reduced transcription rate, as Pol II occupancy was comparable in both strains (Fig. 20). Also the level of Nrd1-HA protein was not changed (data not shown). The most visible difference in Nrd1 association was detected along the body of the gene and upstream of the terminator region, whereas further downstream the Nrd1 level in *cbp80Δ** cells approached the wild-type situation. This observation suggests that in the absence of CBC Nrd1 recruitment to sn/snoRNA genes is delayed.

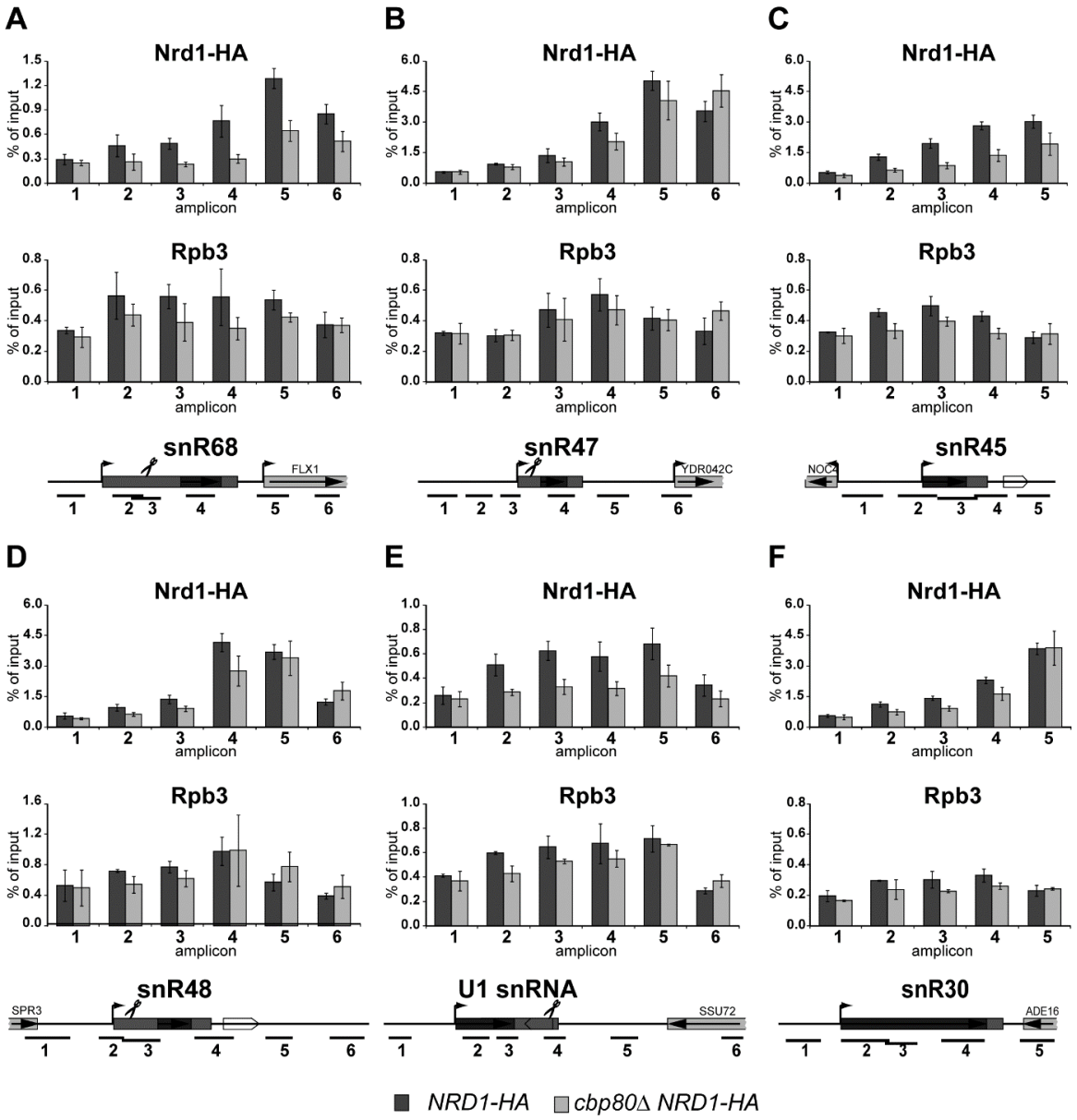

Figure 20. Lack of CBC delays Nrd1 recruitment to sn/snoRNA genes. *(**A-F**) Nrd1 (Nrd1-HA) and Pol II (Rpb3) profiles at different sn/snoRNA genes. qPCR analysis of the ChIP output in NRD1-HA (dark bars) and cbp80Δ NRD1-HA (light bars) strains using anti-HA and anti-Rpb3 antibodies. Description as for Figure 14. The scale on diagrams is set for 6.0 % for most anti-Nrd1-HA IPs and 0.8% for most anti-Rpb3 IPs.*

It was shown recently that CBC affects CTD phosphorylation at mRNA genes through its interaction with Bur1/2 and the Ctk1/2/3 kinase complexes (32, 33). In particular, the level of Ser2-P Pol II was markedly reduced in the absence of CBC. My ChIP analysis using Ser2, Ser5 and Ser7 phospho-specific CTD antibodies did not show significant differences in Ser5-P or Ser7-P CTD at tested sn/snoRNA genes in the *Nrd1-HA cbp80Δ* strain (Fig. 21). However, although Ser2-P is generally low at relatively short sn/snoRNA genes, there was a slight increase in Ser2-P at some snoRNA genes in the absence of CBC, especially at snR47 and snR48 (Fig. 21B and D). What is worth noting, Nrd1 level at these genes was comparable

to the wild type, however, its profile was slightly shifted towards the 3' end. It may suggest some correlation between a delayed Nrd1 recruitment and Ser2-P, although these effects are rather subtle and may not be directly caused by the absence of CBC.

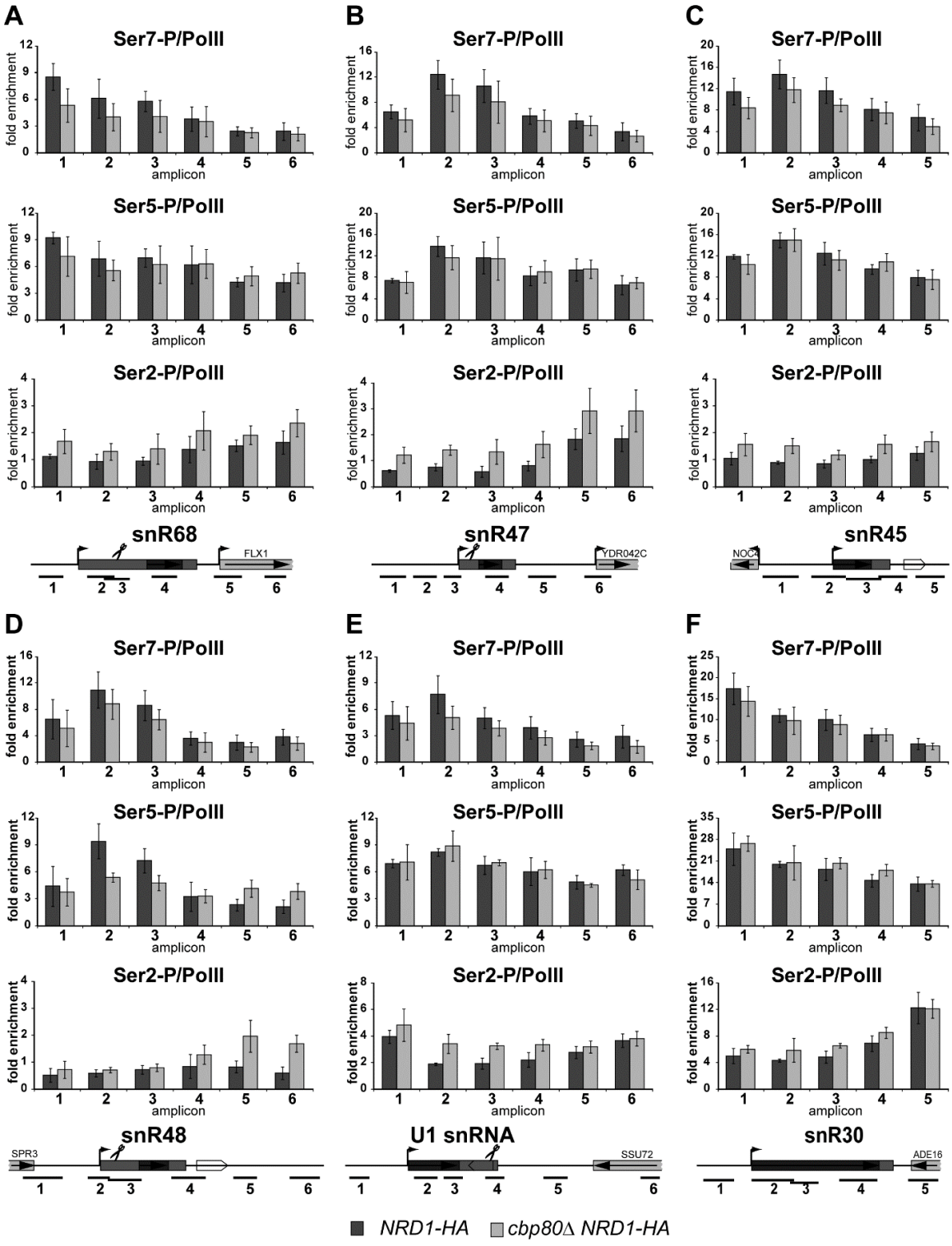

Figure 21. Pol II CTD phosphorylation on sn/snoRNA genes. *(**A-F**) Serine 7 (Ser7-P), Serine 5 (Ser5-P) and Serine 2 (Ser2-P) profiles at different sn/snoRNA genes. qPCR analysis of the ChIP output in NRD1-HA (dark bars) and cbp80Δ NRD1-HA (light bars) strains using anti-Ser7-P (24E12), anti-Ser5-P (3E8) and anti-Ser2-P (3E10) antibodies. Description as for Figure 14.*

3.6 5’ end processing defect causes accumulation of 3’ extended snoRNA precursors

The majority of independently transcribed box C/D snoRNAs undergoes 5’ end processing initiated by endonucleolytic cleavage by Rnt1 that enables further exonucleolytic trimming by exonuclease Rat1 (120). Rnt1 cleavage removes the cap structure associated with CBC. In contrast, 5’ end of snRNA and most box H/ACA snoRNAs corresponds to TSS.

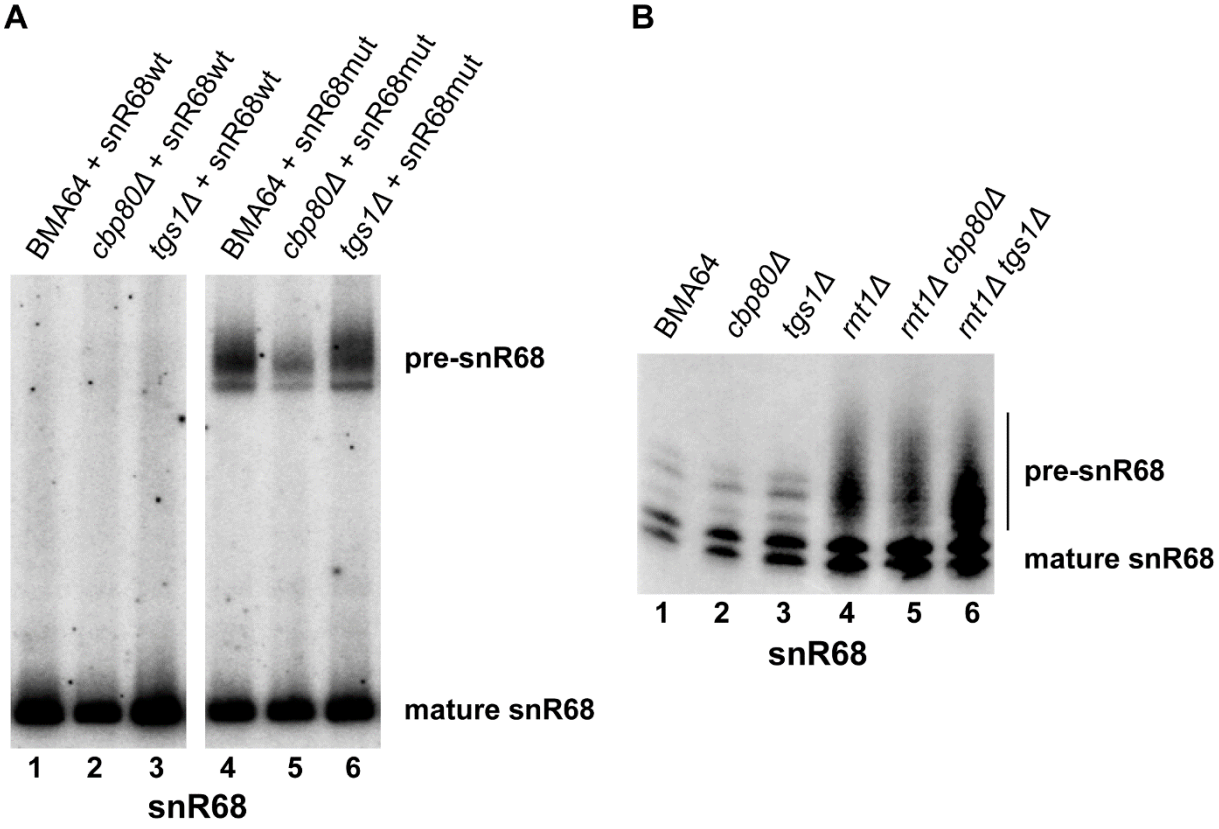

Figure 22. Defective processing of pre-snoRNA 5’ end results in the accumulation of 3’ extended precursors. ***(A)*** *Northern blot analysis of snR68 in wild type (BMA64), cbp80Δ and tgs1Δ strains expressing plasmid-borne wild-type (lanes 1-3) or mutated (lanes 4-6) snR68. The mutated version of snR68 carries two substitutions in the tetraloop recognized by Rnt1 (AGGA→ACAA). Hybridization after RNase H treatment in the presence of an oligo specific for the sequence localized upstream the mature snoRNA. Mature snR68 (genome and plasmid encoded) are visible at the bottom and plasmid-derived pre-snR68 is visible on the top (lanes 4-6).* ***(B)*** *Northern blot analysis of snR68 in wild type (BMA64), cbp80Δ, tgs1Δ, rnt1Δ, rnt1Δ cbp80Δ and rnt1Δ tgs1Δ strains. Hybridization after RNase H treatment in the presence of an oligo specific for the mature snoRNA.*

In this case, their m^7G cap becomes hypermethylated to trimethylguanosine cap (TMG) by the cap specific methyltransferase Tgs1 (108). TMG-capped RNAs do not further bind CBC whose affinity to 2,2,7- trimethylguanosine is much lower than to m^7G (115). Thus, CBC removal from sn/snoRNPs is a common step in maturation of all sn/snoRNAs. Northern blot analysis of three box C/D snoRNA, snR64, snR65 and snR68, showed that 3’ extended snoRNA precursors with highly heterogeneous 3’ ends (3’-pre-snoRNAs) accumulated in the

absence of Rnt1 (Grzechnik, unpublished results). 3' end extensions in *rnt1Δ* cells were even more dispersed than in the case of the *rrp6Δ* mutant, where snoRNA 3' end processing is defective (251, 326, 327, 453).

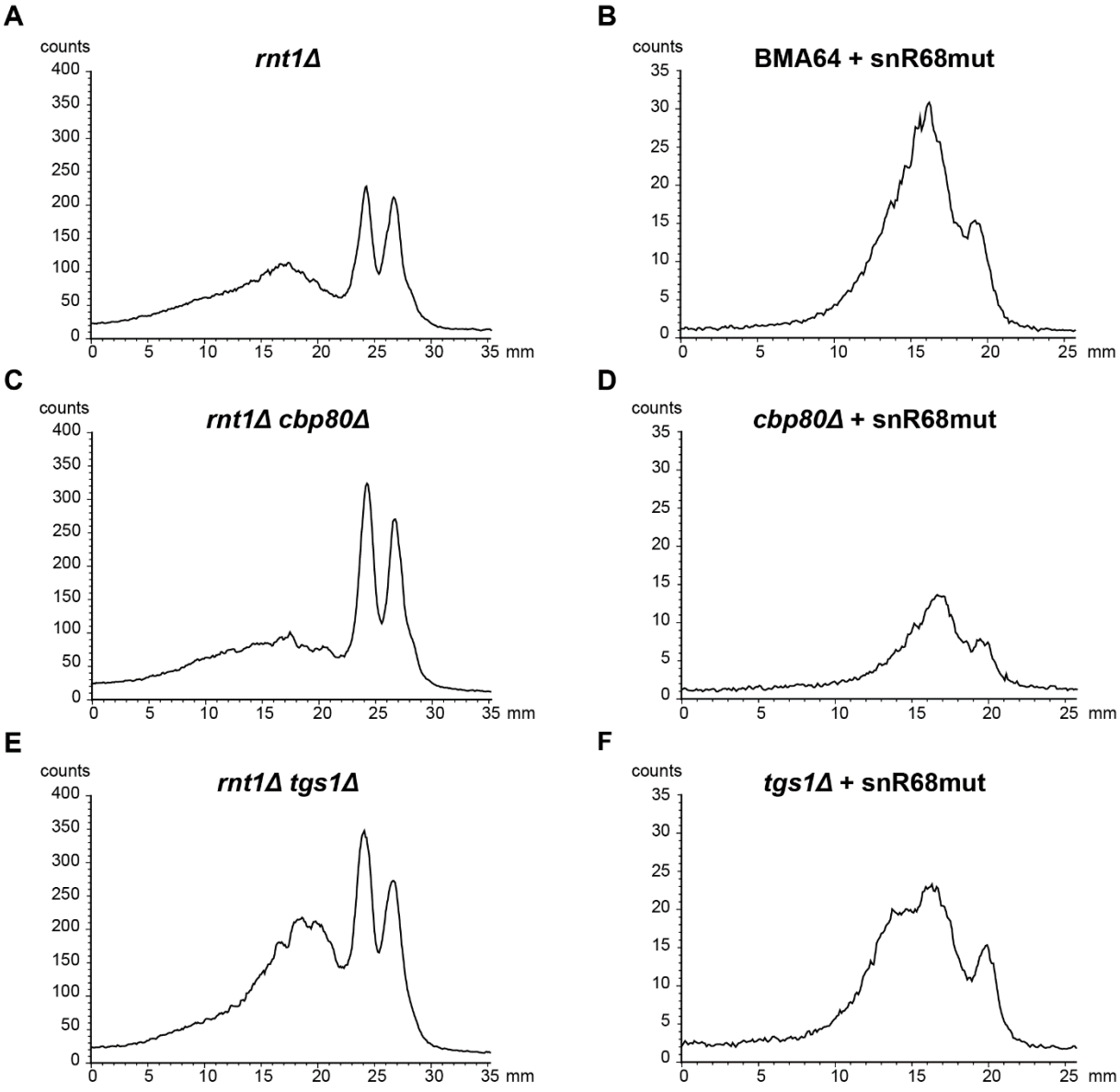

Figure 23. The absence of CBC or Tgs1 contributes to 3' end processing defect. *Densitometric analysis of autoradiographs presented on Figure 22. All bands (mature and pre-snR68) were analysed for rnt1Δ, rnt1Δ cbp80Δ and rnt1Δ tgs1Δ strains (**A**, **C**, **E**) as they were interfering with each other, while only the upper band (pre-snR68) was analysed for BMA64, cbp80Δ, tgs1Δ strains expressing the plasmid-borne snR68mut (**B**, **D**, **F**). For each lane a vertical line was placed manually, the relative density was measured and the diagram was created using the ImageQuant program.*

To further investigate this phenotype I performed similar northern blot analysis in yeast strains expressing plasmid-borne snR68mut carrying mutations in the tetraloop recognized by Rnt1 that disable Rnt1 cleavage from the plasmid. To examine the 3' end more precisely and minimize the background of the wild type snR68, RNA was treated with RNase H in the

presence of oligonucleotide complementary to the 5' end precursor region just upstream of the mature 5' end. Tetraloop mutation caused comparable 3' end processing defect as *RNT1* deletion (Fig. 22A), confirming that endonucleolytic cleavage at the 5' end is important for efficient exonucleolytic trimming of box C/D snoRNA 3' ends. To examine the relationship between Rnt1 cleavage, cap hypermethylation and CBC presence I extended this analysis for strains with deletions of *CBC80* or *TGS1*, either for single mutants expressing plasmid-borne snR68mut or double *rnt1Δ cbc80Δ* and *rnt1Δ tgs1Δ* mutants. In the latter case, oligonucleotide used for RNase H treatment was complementary to the mature snR68 sequence to obtain a better resolution of the diffused extended pre-snR68 population. Interestingly, additional deletion of *CBC80* led to a visibly weaker accumulation of the 3' extended precursor pool in both the *rnt1Δ cbc80Δ* double mutant and for snR68mut in *cbc80Δ* cells, whereas lack of Tgs1 caused a slightly stronger phenotype, namely an increased level of 3'-pre-snoRNAs in the case of the *rnt1Δ tgs1Δ* strain or more extensive 3' end heterogeneity for snR68mut in the *tgs1Δ* mutant (Fig. 22A and B). These observations were substantiated by the densitometric analysis (Fig. 23).

These data suggest that the key element that enables proper 3' end maturation of snoRNAs is the removal of CBC from the 5' end, either by Rnt1 cleavage or by cap modification by Tgs1.

3.7 snoRNA precursors that accumulate in the absence of Rnt1 are oligoadenylated

It was shown before that oligoadenylation is important for the efficient snoRNA 3' end processing (251). Northern blot results described above suggest that in strains where processing of the 5' end is defective, 3' ends of accumulated pre-snoRNAs are oligoadenylated (Grzechnik, unpublished results). To further examine the length and composition of these 3' end extensions I performed circular RT-PCR (cRT-PCR) analysis followed by sequencing for pre-snR68 in *rnt1Δ, cbp80Δ, tgs1Δ, cbp80Δ rnt1Δ* and *tgs1Δ rnt1Δ* strains (Fig. 24-27). Total RNA was treated with RNase H in the presence of oligonucleotide complementary to the 5' region of the precursor between two Rnt1 cleavage sites or Tabaco Acid Pyrophosphatase (TAP) to remove RNA 5' cap structure (Fig. 24A). RT reaction on circularized total RNA was performed with a mature snR68-specific primer to target both mature and precursor snR68 (Fig. 24A). Different combinations of primer pairs were used to amplify the ligated 5'-3' junction, with at least one complementary either to the

3’ or 5’ precursor sequence, to detect only pre-snR68 (Fig. 24A and pre-snR68 scheme on Fig. 25-27).

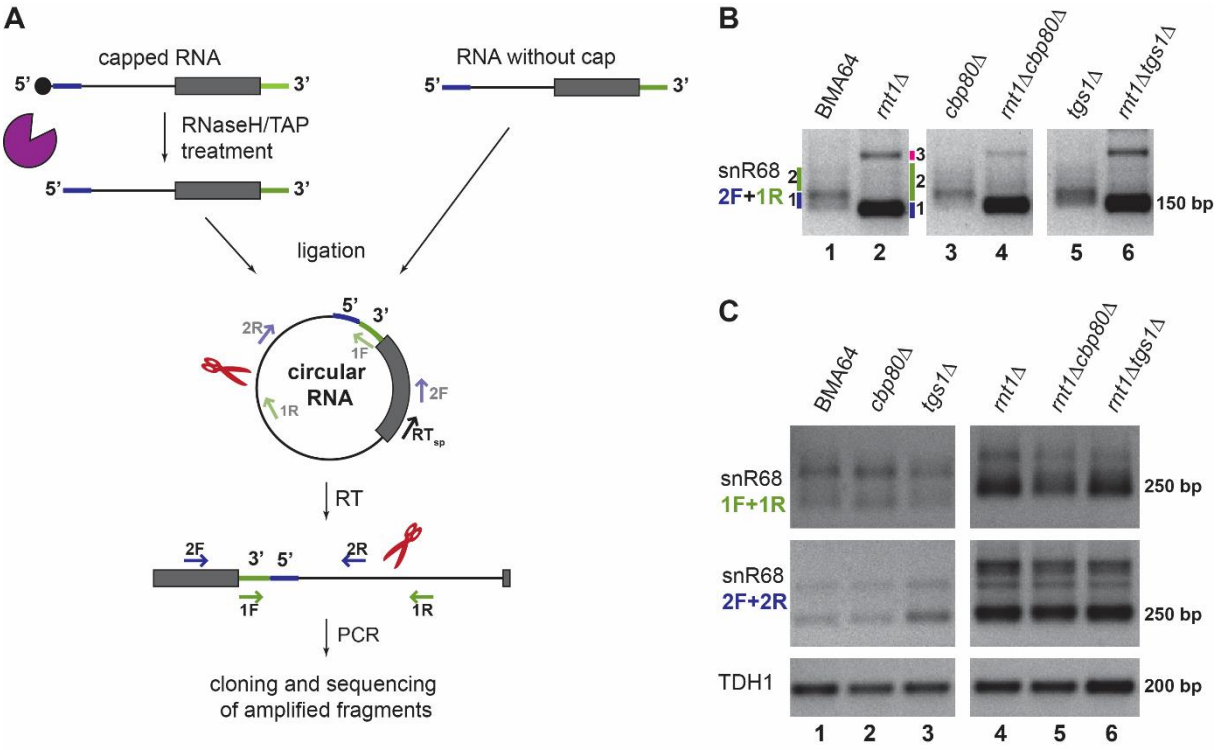

Figure 24. cRT-PCR analysis of pre-snR68 accumulated in the wild type (BMA64), cbp80Δ, tgs1Δ, rnt1Δ, rnt1Δ cbp80Δ and rnt1Δ tgs1Δ strains. *(**A**) Scheme of the cRT-PCR experiment. To remove the cap structure that prevent RNA ligation, total RNA was treated with Tabaco Acid Pyrophosphatase (TAP) or RNase H in the presence of an oligo complementary to the region located within the Rnt1-cleaved hairpin structure. After ligation with T4 RNA ligase, circulated RNAs were reverse transcribed using a primer specific for the mature snR68 (RTsp on the circular RNA diagram). Obtained cDNA was used in PCR or semi-quantitative PCR reactions using 3 different combinations of 4 primers (represented as green and blue arrows): pairs 1F + 1R (green) and 2F + 2R (blue) were used on cDNA obtained from the TAP-treated RNA, pair 2F + 1R on cDNA obtained from the RNaseH-treated RNA. Fragments amplified with primers 1F + 1R (collectively, DNA purified from both bands was pulled together) and 2F + 1R (DNA purified from different bands was treated separately) were cloned and sequenced. Dark rectangle corresponds to mature snR68 sequence, black dot corresponds to cap structure, Rnt1 cleavage site is indicated by scissors. Blue line represents sequenced 5’end, whereas green line represents 3’end. Primers are indicated by black (RT primer), green and blue (PCR primers) arrows. Packman represents TAP/RNaseH. (**B**) Semi-quantitative PCR of the representative experiment performed on the RNaseH-treated RNA using primer pair 2F + 1R. The size of the main product obtained for rnt1Δ, rnt1Δ cbp80Δ, rnt1Δ tgs1Δ strains (lanes 2, 4, 6) is given on the left. Coloured, numbered (1-3) bars near the first panel (lanes 1-2) indicate the fragments purified and sequenced separately. The smeared band in lanes 1, 3, 5 (strains BMA64, cbp80Δ, tgs1Δ) indicated by a green bar (2) consists of 3’ end extended fragments longer than the lower band (blue bar, 1), while sequences of fragments that correspond to the area marked by a green bar (2) in lanes 2, 4, 6 are no longer that the fragments obtained from the main band (bar, 1). The upper band in lanes 2, 4, 6 (indicated by a magenta bar, 3) corresponds to the fragment containing multiple copies of pre-snR68 (most often 2 copies, either identical or different). (**C**) Semi-quantitative PCR of the representative experiment performed on the TAP-treated RNA using primer pairs 1F + 1R and 2F + 2R. The size of main products for rnt1Δ, rnt1Δ cbp80Δ and rnt1Δ tgs1Δ strains (lanes 4-6) is given on the left. Upper bands correspond to the fragment containing multiple copies of pre-snR68. Lower panel shows loading control (TDH1).*

cRT-PCR is a more sensitive technique than northern blot, therefore 3’ extended precursors were detected in all tested strains, but semi-quantitative analysis showed they were

much more abundant in strains lacking Rnt1 (Fig. 24B and C), which is in consistent with northern results. The level of precursors detected with 1F+2R and 2F+2R primer pairs was also elevated in the *tgs1Δ* mutant compared to wild-type or *cbp80Δ* cells, although this effect is more subtle and therefore was not detected by northern analysis (Fig. 24B and C).

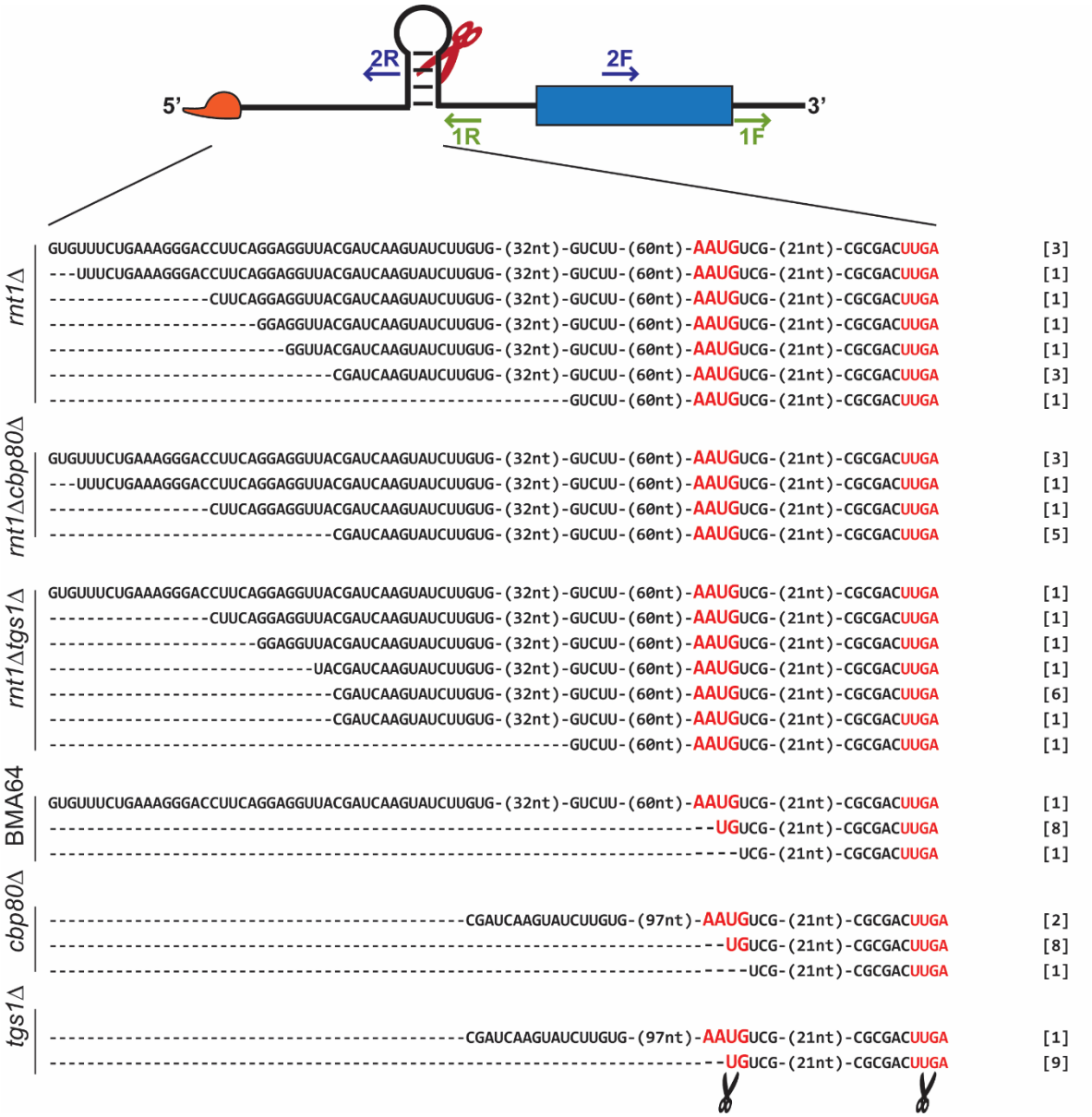

Figure 25. snoRNA precursors that accumulate in the absence of Rnt1 have heterogeneous 5' ends. *Sequences of 5' ends of pre-snR68 in rnt1Δ, rnt1Δ cbp80Δ, rnt1Δ tgs1Δ, cbp80Δ, tgs1Δ and wild-type (BMA64) strains obtained by cRT-PCR analysis. Scheme above sequences represents pre-snR68 organization, the blue rectangle corresponds to mature snR68, red cap represents cap structure, hairpin contains Rnt1-recognized tetraloop and scissors indicate Rnt1 cleavage site. The location of the two pair of primers (1F + 1R and 2F + 2R) is shown as green and blue arrows, respectively. Sequences not shown are denoted by numbers (length) in round brackets; AAUG and UUGA motifs shown in red contain Rnt1 cleavage sites indicated by black scissors (+145 bp and +180 bp from the longest 5'end). Numbers in square brackets on the right represent the number of the same sequences obtained for a particular strain. Location of the sequenced fragment in the precursor is shown by thin lines on the pre-snR68 scheme.*

Sequencing of pre-snR68 5’ and 3’ extensions, together with semi-quantitative analysis, revealed several important points. All snR68 precursors that accumulated in strains lacking Rnt1 were not cleaved at their 5’ ends, whereas 80-90% of 3’ end extended precursors detected in wild-type, *cbp80Δ* and *tgs1Δ* strains were cleaved at the first Rnt1 site (Fig. 25). Interestingly, most of these Rnt1-processed precursors were not further trimmed exonucleolytically. The majority of snR68 3’ extensions in the *rnt1Δ* strain were quite short, composed of 2-5 nt precursor sequence and additional 1-19 A residues (Fig. 26). This pool, amplified with the 1F+2R pair, gave rise to an abundant PCR product that was shorter than in the wild-type strain (Fig. 24B). Similarly short 3’ end extensions (1-5 nt precursor sequence plus 0-5 A residues) in the *rnt1Δ* strain were also observed by cRT-PCR for pre-snR68 using primers specific for the mature sequence (Grzechnik, unpublished results). These results were included in the “RNaseH 2F + 1R” set for BMA64 and *rnt1*Δ strains. Longer 3’ extended snR68 precursors were more common in wild-type, *cbp80Δ* and *tgs1Δ* strains (Fig. 27). These tend to reach the region containing Nrd1/Nab3 binding sites, usually upstream or within the first site, and probably they represent the full-length precursors that are normally present in the cell at a very low level. A few much longer snR68 precursors were detected in *cbp80Δ* and *cbp80Δ rnt1Δ* strains (Fig. 26-27) that reach into the beginning of the next *FLX1* gene. Longer extensions above 24 nt of the precursor sequence were more rare in the *rnt1Δ* strain and were detected mainly by amplification with the primer 1R complementary to the 3’ precursor region. Only a small fraction of these longer 3’ ends were oligoadenylated and their oligo(A) tails were generally shorter (1-7 As) (Fig. 26-27). Unexpectedly, there was no significant difference in the length of 5’ and 3’ extensions between single *rnt1Δ* and double *cbp80Δ rnt1Δ* and *tgs1Δ rnt1Δ* mutants. And finally, this analyses revealed a considerable heterogeneity of precursor 5’ ends that is consistent with previous primer extension results (120) and is probably caused by multiple TSS (Fig. 25).

These data strongly suggest that maturation of both snoRNA termini is synchronized, most likely via CBC bound to the nascent transcript 5’ end until transcription termination, resulting in a delay of processing at the 3’ end. However, even following the release of the transcript, in cases when 5’ end is incompletely or improperly matured, the processing at the other termini is obstructed, leading to the accumulation of 3’ extended precursors.

Figure 26. snoRNA precursors that accumulate in the absence of Rnt1 are oligoadenylated. *Sequences of 3' ends of pre-snR68 in rnt1Δ, rnt1Δ cbp80Δ, rnt1Δ tgs1Δ, cbp80Δ, tgs1Δ and wild-type (BMA64) strains. Two different types of experiments are indicated on the left: shorter sequences were identified using primers 2F + 1R (RNase H treatment), whereas longer and more heterogeneous sequences using primers 1F + 1R (TAP treatment). For the rnt1Δ strain, the first set (RNaseH 2F + 1R) includes initial data obtained by Dr Pawel Grzechnik. Sequence of the 1F primer within pre-snoRNA 3' end is shown in green; in the case of longer 3' ends only the beginning and the end of the sequence is shown, interrupted with a dash. Untemplated oligo(A) tails are shown in magenta, the number of A residues is in subscript (separated by + when more than one different species were found). Asterisk indicates the end of the longest 3' end located within the next ORF. Sequences not shown are denoted by numbers (length) in round brackets. Numbers in square brackets on the left represent the number of the same sequences obtained for a particular strain. Location of the sequenced fragment in the precursor is shown by thin lines on the pre-snR68 scheme (description as for Figure 25).*

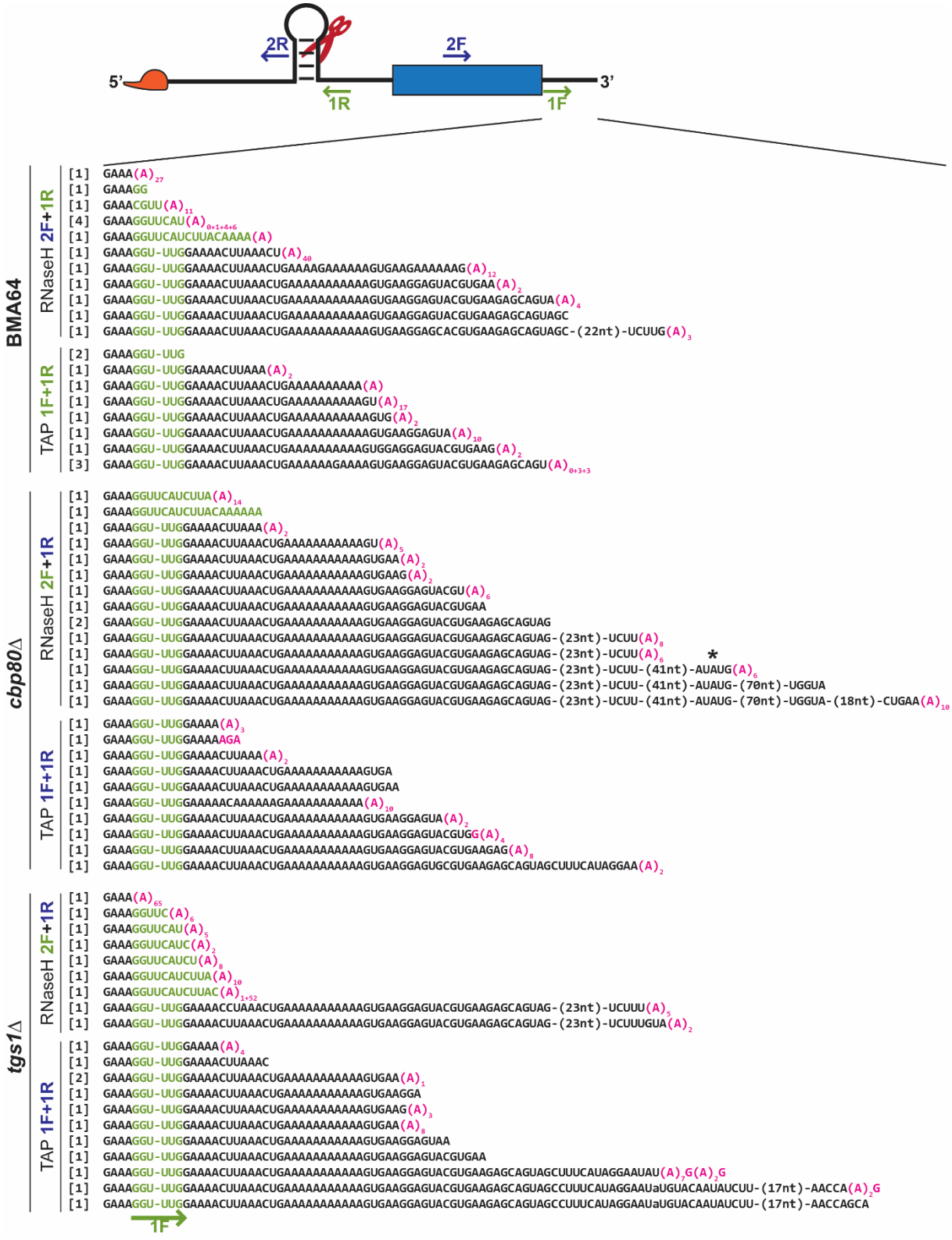

Figure 27. Sequences of 3' ends of pre-snR68 in cbp80Δ, tgs1Δ and wild-type (BMA64) strains. *Description as for Figure 26. For the wild-type strain, the first set (RNaseH 2F + 1R) includes initial data obtained by Dr Pawel Grzechnik.*

3.8 The contribution of the Dcp1/Dcp2 complex to snoRNA processing

Although box C/D pre-snoRNAs accumulate in strains lacking Rnt1 activity, some amount of mature snoRNAs, varying for different species, is still generated (Fig. 22B) (119, 120, 131, 149). An alternative, Rnt1-independent 5' end maturation pathway may involve the removal of the 5' cap structure by the decapping Dcp1/Dcp2 complex followed by trimming by Rat1/Xrn1. Although snoRNA precursors that accumulate in the *rnt1Δ* strain were shown to carry a TMG cap (119, 120, 131, 149), yeast Dcp2 is able to decap TMG-capped RNAs (188).

The level of box C/D pre-snoRNA in strains devoid of the decapping activity has been analysed previously, with some effects detected only for pre-snR68, but not other snoRNAs tested, in *dcp1Δ* and *dcp2Δ* strains and in temperature sensitive *dcp1-2* mutant at the restrictive temperature, especially in the poly(A)$^+$ RNA fraction, suggesting that snR68 precursors were polyadenylated (Grzechnik, unpublished results).

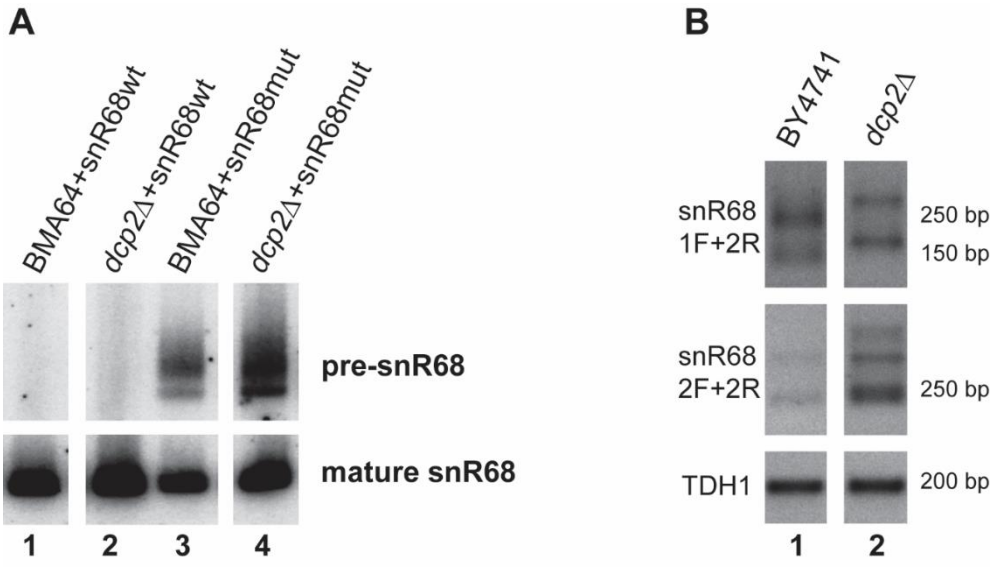

Figure 28. Contribution of the Dcp1/Dcp2 complex to snoRNA processing. *(**A**) Northern blot analysis of snR68 in the wild type (BY4741) and dcp2Δ strains expressing plasmid-borne wild-type (lanes 1-2) or mutated (lanes 3-4) snR68. Description as for Figure 22A. (**B**) Semi-quantitative PCR of representative cRT-PCR experiment performed on TAP-treated RNA using primer pairs 1F + 1R and 2F + 2R for the wild-type (BY4741) and dcp2Δ strains. Description as for Figure 24C.*

The level of snoRNA precursors in decapping deficient strains was much lower than in the *rnt1Δ* strain, which indicates that processing involving cap hydrolysis represents a minor pathway when Rnt1 is active. Attempts to generate double *rnt1Δ dcp1Δ* or *rnt1Δ dcp2Δ* mutants were not successful, suggesting that deletions of *RNT1* and either *DCP1* or *DCP2* may be synthetic lethal. To circumvent this difficulty I used a *dcp2Δ* strain expressing snR68mut with an Rnt1 binding site mutation. The level of unprocessed snR68mut precursors

was greatly enhanced by *DCP2* deletion compared to that in the wild-type strain, showing that lack of the decapping activity adds to the processing phenotype induced by the inhibition of Rnt1 cleavage (Fig. 28A). The accumulation of pre-snR68 in the *dcp2Δ* mutant was also confirmed by the semi-quantitative cRT-PCR analysis (Fig. 28B). Surprisingly, I observed a significant enrichment of precursors that were not cleaved by Rnt1 (primers 2F+2R). Sequencing of cRT-PCR products generated with the 1F+2R pair revealed that approximately 80% of 3' extended precursors in the *dcp2Δ* strain was not processed by Rnt1, compared to only 10% in the isogenic wild-type strain (Fig. 29A). These results suggest that decapping may precede and facilitate Rnt1 processing. Unexpectedly, 5' extensions in *dcp2Δ* cells were on average shorter than those in *rnt1Δ* strains (compare Fig. 25 and 29A), but it may reflect different genetic backgrounds.

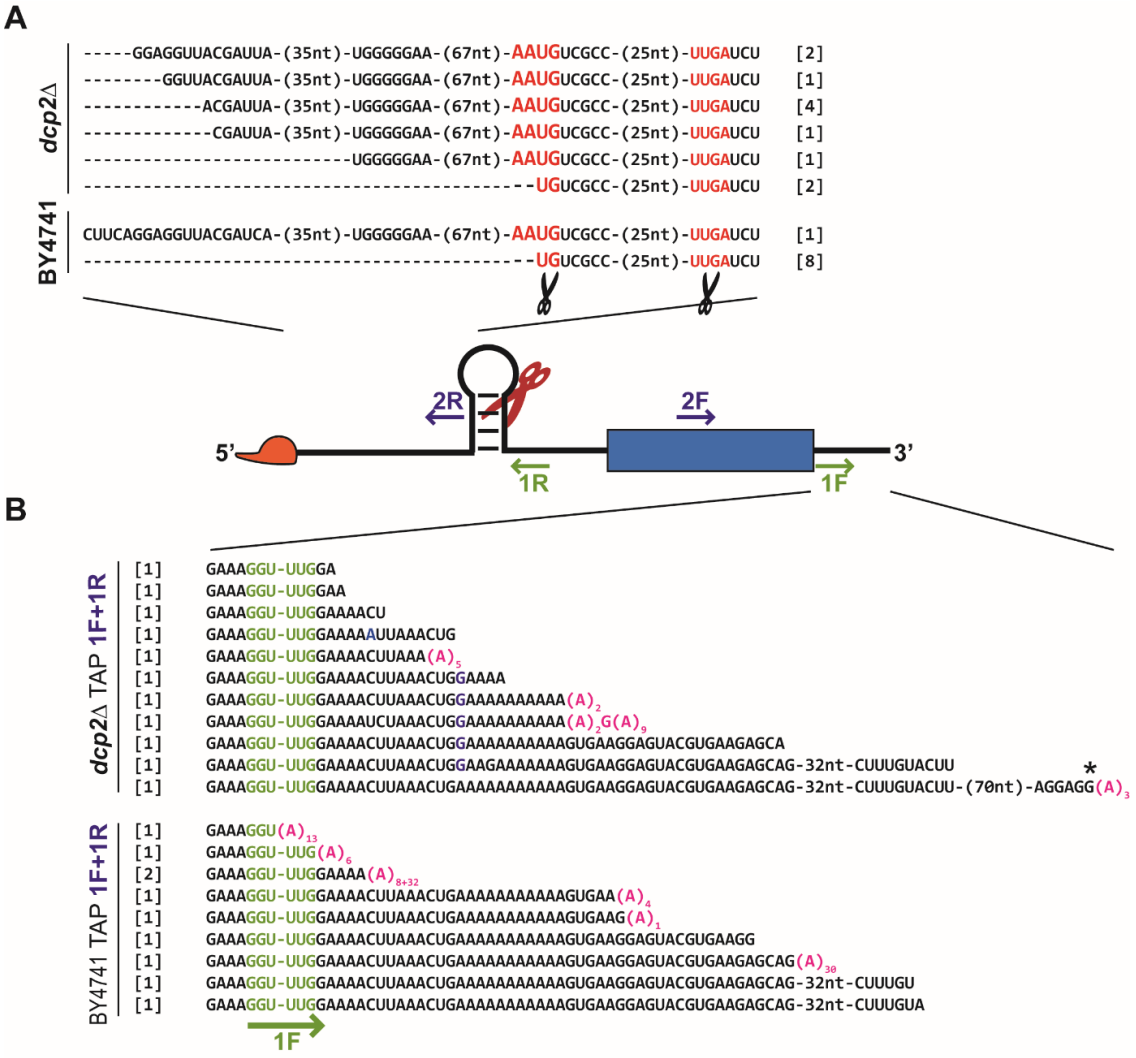

Figure 29. Contribution of the Dcp1/Dcp2 complex to snoRNA processing. *(**A**) Sequences of 5' ends of pre-snR68 in dcp2Δ and wild-type (BY4741) strains. Description as for Figure 25. Rnt1 cleavage sites are indicated by black scissors (+130 bp and +164bp from the longest 5'end). (**B**) Sequences of 3' ends of pre-snR68 in dcp2Δ and wild-type (BY4741) strains. Sequences were obtained using primers 1F + 1R (TAP treatment). Description as for Figure 26.*

3.9 Summary

Multiple roles of m^7G cap and CBC in pre-mRNA processing are well established, whereas their contribution to ncRNA biogenesis was not studied that extensively. Presented data strongly indicate that there is a regulated connection between two termini of pre-sn/snoRNA, since 5' end processing defect results in accumulation of 3' end extended precursors in mutants related to RNA 5' end status (*rnt1Δ, tgs1Δ, cbp80Δ rnt1Δ, tgs1Δ rnt1Δ* and *dcp2Δ*). The presence of cap and CBC on the precursor molecule acts as a signal that its 5' end is still not completed and ready for further processing. A key element of this coordination mechanism involves cap removal (mainly through Rnt1 cleavage) or modification (hyperphosphorylation by Tgs1). CBC directly interacts with ncRNA-specific termination factors, the NNS complex, and remains bound to the nascent transcript throughout the course of its transcription. Early co-transcriptional recruitment of Rnt1 to snoRNAs requires specific stem-loop structure within the precursor, but it appears that endonucleolytic cleavage occurs only at the end of transcription, acting to activate both 5'- and 3' end trimming. This synchronized processing of both sn/snoRNA termini may serve as an additional mechanism regulating error-free and well-timed expression of ncRNAs.

4. DISCUSION

While the role of the 5' cap and CBC in pre-mRNA processing and export of mRNAs and metazoan snRNAs is well established (reviewed in (73, 74)), their function in yeast sn/snoRNA biogenesis remains largely unknown. It has been previously demonstrated that CBC and ncRNA-specific NNS termination factors exist in a multisubunit complex that also contains the exosome (103), but the character and contribution of this complex to sn/snoRNA biogenesis has not been properly explored. Using new affinity resin with attached cap analogue (m^7GpppA-Sepharose) and m^7GTP-Sepharose I showed that the CBC-NNS interaction is direct, independent of RNA and other proteins. CBC is co-transcriptionally recruited to the nascent pre-sn/snoRNA m^7G cap and remains bound until termination takes place. I envisage that during transcription elongation CBC contributes to Nrd1 recruitment to sn/snoRNA genes via its physical connection with the NNS complex. This association links 5' cap and CBC to sn/snoRNA transcription termination and their 3' end processing. Similar relationship was recently described in human cells, where CBC interacts with the ARS2 protein and the CBC-ARS2 complex may be regarded as functionally analogous to the yeast NNS complex (95, 96).

Although 5' end processing of sn/snoRNAs differs between distinct groups, it always involves CBC removal either by combined endo- and exonucleolytic trimming or by cap hypermethylation. My analyses revealed accumulation of 3' extended precursors of box C/D snoRNAs when their 5'end processing is inhibited, mainly in the absence of Rnt1 cleavage. Notably, the additional lack of CBC partially rescues this phenotype, whereas deletion of *DCP2* or *TGS1* has the opposite effect. Together, my data suggest the existence of a quality control mechanism that coordinates the processing at both sn/snoRNA termini, with the CBC-NNS complex being a key factor that prevents premature maturation of either end.

4.1 The CBC-NNS complex coordinates processing of sn/snoRNA 5' and 3' ends

Although the existence of the CBC-NNS complex is well established, its function has not been conclusively demonstrated. Since the NNS complex in involved in transcription termination and nuclear RNA surveillance, it may be presumed that CBC-NNS may contribute to either of these activities. CBC in association with the exosome has been shown to eliminate mRNAs retained in the nucleus due to abnormal physiological conditions in the nuclear mRNA degradation pathway (DRN) (99–101). However, nuclear location of yeast

sn/snoRNAs is absolutely natural, as they are not known to leave the nucleus, so they are not subject to the DRN quality control.

sn/snoRNAs as all Pol II transcripts are capped at the early stage of their synthesis. Although mature sn/snoRNAs do not possess m^7G cap that interacts with CBC, proteomic analyses of sn/snoRNP composition in wild-type and *tgs1Δ* strains revealed CBC association with snoRNP complexes prior to TMG-capping (107). My ChIP analyses demonstrated that CBC is co-transcriptionally recruited to sn/snoRNA genes as soon as cap synthesis is completed, reaching full occupancy about 200 bp downstream of the TSS. This is consistent with a genome-wide CBC association for protein coding genes (33) and in accordance with the ChIP-chip occupancy profiles for sn/snoRNA genes (unpublished data provided by Michael Lidschreiber). CBC presence at the 5' end of pre-sn/snoRNA persists until the completion of their transcription, followed by its rapid removal either by Rnt1 processing or by cap hypermethylation (120). In turn, Nrd1 reaches its full occupancy usually at least 100-200 bp downstream of the CBC peak, so it appears that CBC-NNS complex formation is restricted to a very defined stage of sn/snoRNA biogenesis. This transient and substoichiometric complex assembles most likely co-transcriptionally, probably simultaneously with NNS recruitment, since my data suggest that CBC promotes Nrd1 association with sn/snoRNA genes. The impact of CBC on Nrd1 recruitment is not mediated by changes in CTD phosphorylation status, however, it may be also connected with Pcf11 recruitment to sn/snoRNA terminator. In the case of mRNAs, CBC promotes Ser2-P that stimulates productive Pol II elongation (31–33). However, this effect is not observed for short sn/snoRNA genes, where Ser2-P level is much lower than for protein coding genes.

In the case of transcription termination of protein coding genes, CBC was shown to act as an antiterminator and prevent early termination at weak PASes by interfering with binding of termination factors Rna15 and Pcf11 (86). Both Rna15 and Pcf11 were demonstrated to be involved in NNS-dependent transcription termination. According to recent studies Pcf11, that is recruited downstream of Nrd1-binding sites, cooperates with Nrd1 and facilitates it release (41). It was also suggested that Pcf11 binding enhances Ser2 phosphorylation and promotes Sen1-CTD interaction necessary for its termination activity *in vivo* (37, 292). Slightly elevated Ser2-P level observed in the termination region of some snoRNA genes in the *cbp80*Δ *NRD1-HA* strain can be therefore explained by earlier Pcf11 recruitment and its stimulatory effect on CTD phosphorylation. It would also account for delayed recruitment of

Nrd1 to these genes. The competition between Pcf11 and Nrd1 may also contribute to lower Nrd1 occupancy at sn/snoRNA genes observed in *cbp80*Δ *NRD1-HA* strain.

Three potential function of the CBC-NNS interaction should be considered: targeting to degradation by the exosome as proposed for mRNAs (99–101), transcription termination or 3'end processing activity, similarly to the human CBCA complex (95, 96). Although the first possibility cannot be ruled out and may operate in the case of mRNA and cryptic unstable transcripts (CUTs), my data are more consistent with the second or third explanation.

In humans, the main mechanism involves direct interaction between CBCA and transcription termination or 3' end processing factors (such as CLP1, NELF-E, SLBP15 or FLASH). Thus, CBCA facilitates recruitment of these proteins and stimulates 3' end processing of capped noncoding and coding RNAs (95, 96). During yeast ncRNA transcription CBC may prevent premature binding of Pcf11, at the same time facilitating Nrd1 recruitment to the elongation complex. However, considering that lack of CBC does not cause visible sn/snoRNA termination and 3' end processing defects, this activity is not of great significance, but rather has an auxiliary character. Still, evidently deficient 3' end formation is observed in strains where snoRNA 5' end processing is inefficient, especially *rnt1Δ*. Delayed CBC removal from RNA 5' end in these cells leads to accumulation of 5' and 3' extended box C/D snoRNA precursors. This indicates that CBC-NNS interaction is not indispensable for sn/snoRNA transcription termination, but rather provides quality control mechanism for coordinated 3' and 5' end processing. The important condition of the proposed mechanism is relative instability of the CBC-NNS complex that must disassemble quite rapidly after cap removal or modification (Fig. 30). I envisage that premature processing of both pre-sn/snoRNA ends is precluded by the CBC-NNS complex either by interfering with their RNP assembly or delaying the action of the TRAMP/exosome complex during 3' end formation. Another possibility that cannot be ruled out is that the presence of extended 5' end itself significantly changes RNP conformation and blocks the access to precursors 3' ends, e.g. through base-pairing. Several mechanisms may operate at the same time, since CBC deletion alleviates, but not fully suppresses, accumulation of 3' extended pre-snoRNAs in *rnt1Δ* cells. This may result either from delayed Nrd1 recruitment and consequently also transcription termination or disturbed folding of 5' end extended precursors.

Finally, coordination between processing of both sn/snoRNA termini may involve known interactions between different ncRNA transcription termination and processing factors. Besides Pol II and CBC, the NNS complex, physically interacts with Trf4 and Rrp6, with the

Nrd1 CID domain directly binding Trf4 (103, 273, 280). It is possible that NNS acts within larger complexes, whose dynamic rearrangements are important for synchronizing ncRNA transcription and processing. The connection between Pol II, Nrd1/Nab3 heterodimer (with or without Sen1 helicase) and CBC (forming "Pol II-NNS-CBC complex") would be important to provide full Nrd1 and Nab3 occupancy required for proper ncRNA termination. Following Pcf11 recruitment, Nrd1 dissociates from Pol II CTD and may therefore bind Trf4. Provided that also NNS-CBC interaction is interrupted at this stage, this gives room for Nrd1-Rrp6 association. The connection between NNS, TRAMP and the nuclear exosome, possibly within a larger complex, will then govern either ncRNAs 3' end formation or degradation.

Dissociation of the CBC-NNS complex may therefore play several roles: facilitating Pcf11 recruitment to the elongating polymerase and subsequent association of Sen1 and activation of its transcription termination activity; enabling interaction between NNS, TRAMP and the exosome that directs sn/snoRNA 3' end processing or degradation; allowing 5' end processing synchronized with CBC removal. However, these events are interdependent and it is difficult to discern their timing or mutual impact.

4.2 Rnt1, Tgs1 and Dcp1/Dcp2 synchronize processing of sn/snoRNA 5' and 3' ends

It is well established that pre-mRNA maturation (capping, splicing and 3' end formation) takes place co-transcriptionally. The timing of sn/snoRNA processing and its association with transcription is less obvious, though it was shown that efficient and correct 3' end formation of box C/D snoRNAs requires Nop1 snoRNP core factor (316). The assembly of box H/ACA snoRNPs also occurs co-transcriptionally at an early stage of biogenesis, beginning with the recruitment of Cbf5p and RNA binding protein Naf1 (321, 433). Components that specifically assemble on nascent transcripts to form functional RNPs probably impact their transcription and processing, but coordination of sn/snoRNA processing with transcription and RNP assembly has not been extensively studied. Although 3' end trimming must occur post-transcriptionally, 5' end cap hypermethylation or Rnt1 cleavage may theoretically take place concurrently with RNA synthesis. Co-transcriptional Rnt1 processing was shown for pre-rRNA, which primary transcript is the most abundant Rnt1 substrate in the cell but is hardly detectable *in vivo* (132). I showed that Rnt1 is co-transcriptionally recruited to sn/snoRNA genes but reaches full occupancy near the termination region. Consistently, CBC association with sn/snoRNA genes also persists until the termination region. Lack of Rnt1 activity causes accumulation of 3' extended, oligoadenylated box C/D snoRNA precursors. In wild-type cells

pre-snoRNAs are also present albeit at a very low level and their both termini are also extended, with 5' ends beginning from the Rnt1 cleavage site and not processed exonucleolytically. This data indicates that maturation of both snoRNA termini is synchronized. The cleavage by Rnt1 is rapid and efficient, since almost all precursors detected in strains with active Rnt1 were cleaved at 5' end. In addition, *in vitro* cleavage of Rnt1 substrates occurs extremely fast, a visible fraction of cleavage products can be observed within few seconds (data not shown). Also meta-analysis of native elongating transcript sequencing (NET-seq) data (454) strongly suggests that Rnt1 cleavage of pre-snoRNAs is co-transcriptional (Grzechnik, unpublished). However, Rnt1 and CBC ChIP results indicate that snoRNA 5' end processing is delayed until transcription termination. Rapid Rnt1 cleavage provides entry site for Rat1/Xrn1 to unprotected 5' ends and at the same time activates 3' end processing by removing CBC from pre-snoRNP. Exonucleolytic trimming of both ends is slower than Rnt1 processing, therefore a small subset of extended precursors can be detected.

Rnt1 was also shown to interact with the NNS complex (36, 103, 455) and this may additionally affect sn/snoRNA transcription termination or 3' end processing. Although our ChIP analysis revealed that Rnt1 recruitment to snoRNA genes is not dependent on Sen1, Rnt1 association profiles at snoRNA genes resemble Nrd1 profiles, suggesting that also Rnt1-NNS connection is established co-transcriptionally. Rnt1 may therefore coordinate 5' and 3' end processing in two ways: by removing CBC through the 5' end cleavage or via the interaction with the NNS complex. The second connection could act as the "NNS complex sensor", preventing the cleavage until the NNS complex reaches its optimal occupancy. It is noteworthy that Rnt1 cleavage site at the 5' end of Pol III transcribed pre-snR52 is not strictly required for its 5' processing (120), which additionally underlines the role of Rnt1 cleavage in CBC removal.

Similar, but much lower accumulation of 5' and 3' extended precursors is observed when cap hypermethylation or decapping are disrupted. It was shown before that box C/D snoRNA precursors that accumulates in *rnt1Δ* strain are TMG capped, that suggests rapid cap hypermethylation following pre-snoRNA transcription, but this modification does not additionally stabilize pre-snoRNAs (119, 120). I showed that in the absence of Tgs1 snR68 precursors that are not processed by Rnt1 are present at a very low level. This confirms rapid cap modification that at least partially can occur prior to Rnt1 cleavage. Although Rnt1 processing is hardly affected by the absence of Tgs1, cap hypermethylation may facilitate this process. I envisage that the role of cap hypermethylation is not pre-snoRNA stabilization, but

rather CBC removal from the 5' end and probably also from the complex with NNS. Additional studies are necessary to establish the role of Tgs1 in this process, especially in the case of snRNA and box H/ACA snoRNA. Precursors of these ncRNAs are not 5' extended and do not require 5' end processing, therefore cap hypermethylation is the only step in their maturation that removes CBC from their 5' ends.

The role of the decapping complex Dcp1/Dcp2 in snoRNA processing has been suggested (120), but has never been demonstrated. Most box C/D snoRNA cleaved by Rnt1 at their 5' end do not entirely depend on Rnt1 for their maturation ((120), this study), therefore an alternative pathway that enables exonucleolytic 5' end processing must exist. It was shown in my laboratory that pre-snR68 accumulates at a very low level in mutants with disrupted decapping activity (Grzechnik, Matuszek unpublished results). Moreover, my results show that *DCP2* deletion causes greater accumulation of pre-snR68 with mutated Rnt1 recognized tetraloop. Surprisingly, most of the precursors present in *dcp2Δ* strain was not cleaved by Rnt1. Collectively, this data demonstrate the role of Dcp1/Dcp2 in box C/D snoRNA maturation not only as a part of an alternative 5' end processing pathway, but also in promoting Rnt1 cleavage. Since processing of other box C/D snoRNAs was not affected in Dcp1/Dcp2 deficient strains (Grzechnik, Matuszek, unpublished results), this suggests that additional decapping enzymes may be involved in this mechanism. Besides Dcp1/Dcp2 also Rai1, Dxo1 and Nudt3 have a decapping activity, although first two were shown to preferably hydrolyse unmethylated cap substrates (80, 161, 189, 196). It will be worthwhile to test the potential role of these proteins in the alternative pathway generating mature snoRNA 5' ends.

4.3 The model of pre-sn/snoRNA maturation

It was suggested that 3' and 5' extensions in box C/D snoRNA precursors provide a way to prevent their degradation before completing snoRNP assembly (120, 429). snoRNP formation, involving RNA folding and snoRNP protein binding, begins co-transcriptionally and involve additional auxiliary factors (316, 321, 433, 436). Mature ends of box C/D snoRNAs form a stem structure, so their final folding occurs after 3' end has been transcribed and released from the Pol II complex. It is quite likely that complete snoRNA assembly is achieved after transcription termination, but at least some factors associate during transcription. For example, Nop1 is recruited co-transcriptionally to nascent box C/D pre-snoRNAs and this is probably mediated by Ref2. It is also possible that binding of Snu13 and Nop58 precedes Nop1 association (316, 438).

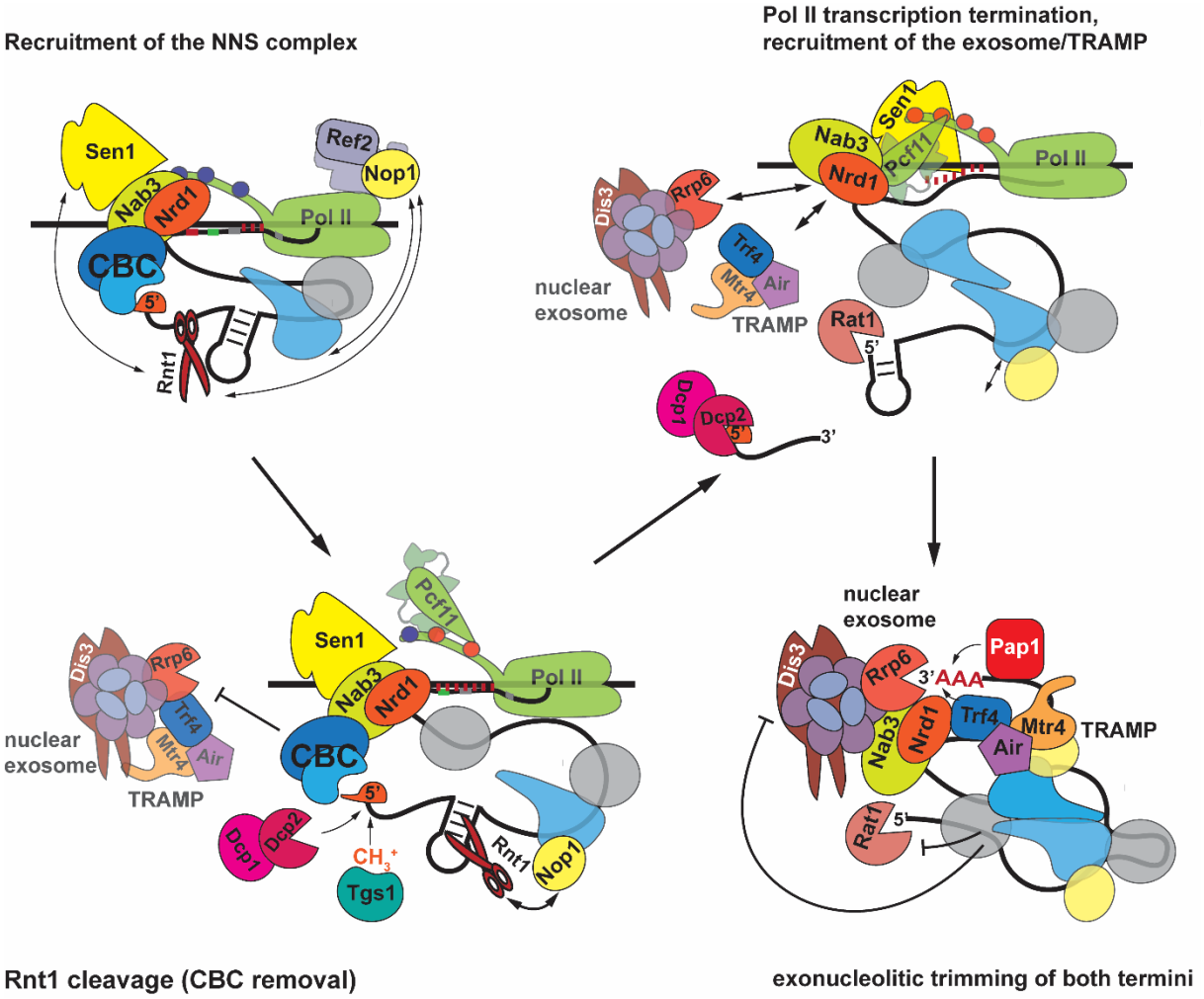

Figure 30. Role of CBC and Rnt1 in box C/D pre-snoRNA processing. *Working model based on presented data and recent publications. Interaction of CBC-NNS complexes is established during snoRNA transcription. snoRNP proteins are recruited co-transcriptionally, beginning with Snu13 and Nop58 in the case of box C/D pre-snoRNAs. Rnt1 is recruited to the tetraloop, reaching full occupancy at the late phase of transcription. Endonucleolytic cleavage of the pre-snoRNA 5' end releases CBC from its interaction with the NNS. NNS complex changes its interacting partners, recruiting the nuclear exosome (possibly through the interaction with Rrp6 cofactor) and TRAMP complex (interaction with Trf4) and after transcript release thanks to the action of helicase Sen1, pre-snoRNA 3' end is oligoadenylated by Pap1 and TRAMP, which facilitates trimming by the exosome. Pre-snoRNA 5' end lacking of the cap structure is trimmed by Rat1/Xrn1. At the same time snoRNP proteins are recruited and the RNA component is folded into its mature structure, which both naturally hinder progression of exonucleases. Interactions between proteins are indicated either by their physical contact or double-arrow lines. Lines with arrowheads indicate catalysed reaction, lines with bar heads indicate inhibition. Red bars are schematic representation of RNA-DNA hybrids. Detailed description in the text.*

Although processing of the 5' end may begin co-transcriptionally, my data suggest that it is rather delayed until transcription termination. Moreover, maturation of both sn/snoRNA termini is coupled and synchronized by the CBC-NNS complex and Rnt1 (Fig. 30). CBC-NNS interaction prevents premature degradation of the 5' end and although Rnt1 is co-transcriptionally recruited to the nascent pre-snoRNP, it probably does not cleave the precursor until the very end of its transcription. This is supported by indications that Rnt1

activity is facilitated by interactions with Sen1 and Nop1 that associate with precursors at the late stage during transcription.

After cap removal (Rnt1 cleavage or decapping) CBC no longer binds the nascent transcript and probably is also released from its interaction with NNS. This enables further 5' end exonucleolytic trimming in the case of most box C/D pre-snoRNAs and at the same time permits NNS to recruit 3' end processing factors, exosome and TRAMP complexes, via direct binding of Rrp6 and Trf4, which activates 3' end processing (Fig. 30). snoRNP proper folding and association of the remaining core proteins create natural boundaries that stop exonucleolytic digestion at the right place. This provides a quality control mechanism ensuring that snoRNAs are properly assembled and have correct, well-defined mature 5' and 3' termini.

5. MATERIALS AND METHODS

5.1 Materials

5.1.1 Yeast Strains

Name	Genotype	Ref or source
BY4741	*MATa his3Δ1, leu2Δ0, met15Δ0, ura3Δ0*	Euroscarf
BMA64	*MATa, ura3-1, ade2-1, his3-11,5, trp1Δ, leu2-3,112, can1-100*	(456)
rnt1Δ	as BMA64 but *RNT1::TRP1*	(131)
cbp80Δ	as BMA64 but *CBP80::HIS3*	This study
tgs1Δ	as BMA64 but TGS1:: *HIS3*	This study
cbp80Δ rnt1Δ	As *rnt1Δ* but *CBP80::HIS3*	This study
tgs1Δ rnt1Δ	As *rnt1Δ* but *TGS1::HIS3*	This study
rrp6Δ	as BY4741 but *RRP6::kanMX4*	Euroscarf
dcp2Δ	*leu2-3112 his4-539 lys2-201 trp1 ura3-52 dcp2::TRP1*	(201)
BMA64+ pRS415-snR68WT	as *BMA64* but *pRS415-snR68WT*	This study
BMA64+ pRS415-snR68mut	as *BMA63* but *pRS415-snR68mut*	This study
cbp80Δ+ pRS415-snR68WT	as *cbp80Δ* but *pRS415-snR68WT*	This study
cbp80Δ+ pRS415-snR68mut	as *cbp80Δ* but *pRS415-snR68mut*	This study
tgs1Δ+ pRS415-snR68WT	as *tgs1Δ* but *pRS415-snR68WT*	This study
tgs1Δ+ pRS415-snR68mut	as *tgs21Δ* but *pRS415-snR68mut*	This study
dcp2Δ+ pRS415-snR68WT	as *dcp2Δ* but *pRS415-snR68WT*	This study
dcp2Δ+ pRS415-snR68mut	as *dcp2Δ* but *pRS415-snR68mut*	This study
GAL::CBP20 (Y258)	*MATa, pep4-3, his4-580, ura3-52, leu2-3, 112, pBG1805-GAL1::CBP20-His$_6$/HA/ZZ* (*URA3, 2μ*)	(450)
GAL::eIF4E (Y258)	*MATa, pep4-3, his4-580, ura3-52, leu2-3, 112, pBG1805-GAL1::CDC33-His$_6$/HA/ZZ* (*URA3, 2μ*)	(450)
GAL::CBP20 GAL::CBP80 (Y258)	as *GAL::CBP20* but + *p425GAL1-CBP80-FLAG-2×Strep-tag II*	MSc D. Adamska
Cbp20-Myc	as *BY4741* but *CBP20-myc13::HIS3*	(86)
Cbp80-Myc	as *BY4741* but *CBP80-myc13::HIS3*	(86)
Rad30-Myc	*his3-Δ200, leu2-3,2-112, lys2-801, trp1-1, ura3-52, RAD30-9myc::klTRP1*	(457)
Rnt1-Myc	as *BY4741* but *RNT1-myc13::NAT*	This study
snr68Δ	as *BY4741* but *SNR68::HIS3*	This study
snr68Δ + pRS415-snR68mut	as *snr68Δ* but *pRS415-snR68mut*	This study
Rnt1-Myc snr68Δ+ pRS415-snR68WT	as *snr68Δ* but *RNT1-myc13::NAT, pRS415-snR68WT*	This study
Rnt1-Myc snr68Δ+ pRS415-*snR68mut*	as *snr68Δ* but *RNT1-myc13::NAT, pRS415-snR68mut*	This study
sen1-R302W	as *BY4741* but *sen1-R302W*	(37)
sen1-K128E	as *BY4741* but *sen1-K128E*	(37)
Rnt1-Myc sen1-R302W	as *sen1-R302W* but *RNT1-myc13::NAT*	This study

Rnt1-Myc sen1-K128E	as *sen1-R302W* but *RNT1-myc13::NAT*	This study
Nrd1-HA	as *BY4741* but *NRD1-HA3::NAT*	This study
*cbp80Δ**	*MATa, ade2, ade3, his3, leu2-3, 112 rp1 ura3 CBP80::TRP1*	(84)
cbp80Δ Nrd1-HA	as *cbp80Δ** but *CBP80::TRP1; NRD1-HA3::NAT*	This study

5.1.2 Bacterial strains

E. coli BL21(DE3): *fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*
E. coli Rosetta (DE3): *F- ompT hsdSB(rB- mB-) gal dcm (DE3) pLysSRARE (CamR)*

5.1.3 Plasmids

Plasmid	**Reference or source**
pST39-*Nrd1-His$_9$-Nab3-MBP*	(274)
p425GAL1	(458)
p425GAL1-*CBP80-Flag-StrepTag II*	MSc D. Adamska
pRS415-snR68WT	This study
pRS414-snR68mut	This study
pFA6a-13Myc-natMX6	(459)
pFA6a-3HA-natMX6	(459)
pJET1.2	Fermentas

5.1.4 Oligonucleotides and probes

Primer no	**name**	**sequence**
strains construction		
W332	3HAnrd1FA	GAACATAGGAAAAAACAGAAATTATATATAGAGGTAGA TTAGTTTTATGTGAATTCGAGCTCGTTTAAAC
W333	5HAnrd1FA	ATTCTTTGATGAATATGCTTAACCAACAGCAGCAGCAAC AACAACAAAGCCGGATCCCCGGGTTAATTAA
W336	FNrdHA5v	CCTCGTTAGCATGACTCCC
W337	RNrdHA5v	GTACCGGGTTTGTTGCTGC
W568	Rnt1F2	GTGTTAAAAGATCCCTCACAAAAGAATAAGAAAAGAAA ATTCTCAGATACAAGCCGGATCCCCGGGTTAATTAA
W569	Rnt1R1	TTTATGAATGACATATATGCGATAATTATATGGCTAAAG AAAATCAATGCGAATTCGAGCTCGTTTAAAC
W570	Rnt1_ver_F	AGTGCAGAGTAGGTGATGGAAC
W571	Rnt1_ver_R	AGACGCGTTGCCATTACGC
ChIP		
W443	snR45F1	ATCTCTCTTCATGGCAATAGAGGATG
W444	snR45R1	CACTTCTCTACGGGTAATCCTCG
W445	snR45F2	AGGTATAAAAAGCGAAACACTCGGTAC
W446	snR45R2	TCCGAGAAGAATTGTTCGATTTGAAACG
W447	snR45F3	TCGGAGCGATCTGAGGTTTTAATGG
W448	snR45R3	GTGTACAGATGAGATGACTACTCCCAAG
W449	snR45F4	ACTTGGGAGTAGTCATCTCATCTGTAC
W450	snR45R4	CGGCAAGTTTCCCTGGATGTC
W451	snR45F5	GGTCTCTGTGGAAACCAGTGC
W452	snR45R5	GTAGTTGATGCCCGAAGGTGC
W453	snR82F1	TGACAGCGTTATTTGAGAATAGACCTAAG

W454	snR82R1	GTTGAAATGTGTTGAAGAGCCATGAT
W455	snR82F2	ATCATGGCTCTTCAACACATTTCAAC
W456	snR82R2	TCTTGAGCTTATTTCTTAATGCGTGGG
W457	snR82F3	ATCCCGCCTTTTTTCGCTAACAATG
W458	snR82R3	AGGTATCGGTCAGTATCACGCTC
W459	snR82F4	GACAGGACACATAATAGCAAGGCAAG
W460	snR82R4	TTGTCGTCACTTGATCCCAGTTG
W462	snR47F1	ACCTTTCGAGCATCTCTTACAGCAATGG
W463	snR47R1	TGTTCGGAGACAGTTCTGATGCCTAG
W464	snR47F2	CCCTAGAAGAAATACCCGAAGATGTAAG
W465	snR47R2	ACCGTATGGAAGACGTAGAGTGGATG
W466	snR47F3	AAGGCTTCAGCTCCATATC
W467	snR47R3	CCTTTCTCCTACTTTGCTCAG
W468	snR47F4	ACAACAACATGAATTTCTTCGTCCG
W469	snR47R4	CAGCAAGAATGACGCGAAA
W470	snR47F5	TTTCTGTTTCTGTTTCGCGTCGG
W471	snR47R5	TCCCTGTTATCCGCCTTTCTTCTTGG
W472	snR47F6	CAGCTAACAAACGACTAGGTCTC
W473	snR47R6	CCTTGAAAAGTAGAAAGGGTAG
W476	snR48F1	ACAGCCATTCCACAGATAGTGCTAC
W477	snR48R1	TTGCTACTTATGGTAAAACTGCCGC
W478	snR48F2	ATACCTTTGTCCGCTGTTATCATACC
W479	snR48R2	GTTACCGTACCACTCTTGCTTATG
W480	snR48F3	GGTACGGTAACCACAAGGCAATG
W481	snR48R3	GATAAACGAGACGGATGAAAGGAG
W484	snR48F4	ACTATGATTAAACAGACCGAGGGAG
W485	snR48R4	ACGACGCGGACGAAGAGAAAG
W486	snR48F5	TTCGTTACCCGGTGTAGTCACG
W487	snR48R5	CGATAACTGGCTCTGCTCTTTCCTG
W488	snR48F6	ACCGATAACAGTCAGCAACCTCAAG
W489	snR48R6	AATAATACGACCACAGCAGAGGCG
W492	snR30F1	ATGTCCTTCTAGCGTAGTAACC
W493	snR30R1	GTTAAATGCACGACGATGAGAG
W494	snR30F2	ACCATAGTCTCGTGCTAGTTC
W495	snR30R2	CTTATGTGATGCCGTTGTCC
W496	snR30F3n	TCGGACAACGGCATCACATAAG
W497	snR30R3n	ATACACATCGTTAAGGCAACAGC
W498	snR30F4	AAGTAGGACGCATGATCTTG
W499	snR30R4	GGACCAACTAGGGTCATTTC
W500	snR30F5	TCTCATAAAATGTACCCTACTTAGTGG
W501	snR30R5	TGTAAAGTTCATCACGGCTCC
W540	U1F1	TGGTCTTTTTGCTTCCTGGAACTC
W541	U1R1	TATGATGGCAGCGTAGATCACCG
W542	U1F2	GCTGCCATCATAGGGCTCATTG
W543	U1R2	CCAATGGAAAACTCGTCGCCTTAG
W544	U1F3	AGAAGTCCTACTGATCAAACATGCG
W545	U1R3	CGCCGTATGTGTGTGTGACCA
W546	U1F4	CCTTGTTTCAATCATTGGTTAATCCCTTG
W547	U1R4	TGGTGTCAAACTTCTCCAGGCAG
W548	U1F5	AGTTTATGATGAAGGTATGGCTGTTGAG
W549	U1R5	ACGGGAACGAGCAAAGTTGAGAC
W550	U1F6	CTACATTGAGACAAGACATTGTCCAGC

W551	U1R6	ACGTTTCTTGTTGCGGCTCTC
W552	snR41_70_51F1	GTCAACAAGCAAGTATCCCGCAAG
W553	snR41_70_51R1	TCCAGTTCGGGCTAAAGAGTCG
W554	snR41_70_51F2	CTTCCTTTGTTTGCTTCTTCGTCTTC
W555	snR41_70_51R2	ACCTACTGGTCGCTTAATCTGCCTC
W556	snR41_70_51F3	TCTTTCTGTGTGACTCCCCTATG
W557	snR41_70_51R3	ACCCCTCATACAAGGCTGAAAC
W558	snR41_70_51F4	AAGGGTCGCCGTCTACTCTC
W559	snR41_70_51R4	TGTCTTGTGACCAATCATCAATTCTCC
W560	snR41_70_51F5	TGATAAAAGAGACTGTTGCGGACTAG
W561	snR41_70_51R5	TTGTAGTCATCAATTAGCCCCTTCG
W562	snR41_70_51F6	AGGAGAAAGGGTTTTGAACACATCG
W563	snR41_70_51R6	TGTGATACCGCCAAGCAAGTAAAC
W592	snR68F1	AGTCAGCATTCAACAACGTGAGAAG
W593	snR68R1	AGAGCATGACCGTGTATGGTGAC
W594	snR68F2	AGGGACCTTCAGGAGGTTACG
W595	snR68R2n	GTACAAAGCTATGACCAACCGTTCC
W596	snR68F3	TTGAAAAGAGTTAATTATGGAACGGTTGG
W597	snR68R3	TCGTTTGATAGCAGTAAGTCGCAG
W598	snR68F4	TTTACTGCGTTATCGTATTGACGG
W599	snR68R4	AGCAAATCTGTTAAGAGTCAATTTCCTCG
W600	snR68F5	AACCAGCACGCAGCAGGAAG
W601	snR68R5	TGGATGCACTACCAATGTGGTG
W602	snR68F6	TCGGCGAATAGTGGACGTAGTG
W603	snR68R6	GAAATAAACTCCCCATGCAATGGC
	U5F1	GTTGCTGAATTTCCTGACTCTACGG
	U5R1	CTCTACCCATTATCCCTTTGTCATCG
	U5F2	AATGGCGGAGGGAGGTCAAC
	U5R2	ACCTGTTTCTATGGAGACAACACCC
	U5F3	TGTCCGTTACTGTGGGCTTGC
	U5R3	GTTCCATGGACTCATGAATCAAATTTGTAG
	U5F4	TGGAGCAGAAATAAGACGATCACAGG
	U5R4	AGTTTAGTCTCAAATGTATCGAACGCG
	U5F5	ACTCTTGTTTCATCGGTTGTTGCC
	U5R5	AGGGGAAGCACTATTTCTCACTTG
	U5F6	GAACCTTCTATGCAGGATTTTGACCC
	U5R6	TCTTTTCTGTGGCATTGCTGTCTGAG
	snR52F1	ACAACTCTAGATTTTGTAGTGCCCTC
	snR52R1	ACTTTCACTTCTACAGCGTTTGACC
	snR52F2	TCTAGATTTTGTAGTGCCCTCTTGG
	snR52R2	CACCAACTTTCACTTCTACAGCGT
	snR52F3	ACGCTGTAGAAGTGAAAGTTGGTG
	snR52R3	TCAGAGATTGTTCACGCTAATGTCATTC
	snR52F4	TGACATTAGCGTGAACAATCTCTGATAC
	snR52R4	TGCGTTCCATACTGTCAGAGGTG
	snR52F5	CTTGCTAGTTTTGCGTTGCCCT
	snR52R5	AGTTTTTGATGTACTGTATTCCTTGCG
hybridization probes		
W076	snR68-2	AAGAGTCAATTTCCTCGGTA
W271	snR68	GCTTTCAGATACTATCTAGC

RNase H treatment		
	snR68son	CCCCGTCAATACGATAACGC
W315	cRTPCR1R	AATTTATCGTTTGATAGCAG
cRT-PCR		
	68Hlig	GTATCCTTACAAACATGACG
W091	RTsp	GATAACGCAGTAAAATAAATG
	cRTPCR1F	GGTTCATCTTACAAAAAAATTG
W315	cRTPCR 1R	AATTTATCGTTTGATAGCAG
W092	cRTPCR 2F	GTACAGTCTGTTTTATAATC
	cRTPCR 2R	CCAACCGTTCCATAATTA
	TDH1F	GGTATGGCTTTCAGAGTCCCA
	TDH1R	AGACAACGGCATCTTCGGTG
	pJET1.2F	CGACTCACTATAGGGAGAGCGGC
	pJET1.2R	AAGAACATCGATTTTCCATGGCAG

5.2 Methods

Reagents used were purchased from Sigma-Aldrich unless stated otherwise.

5.2.1 Basic techniques used in molecular biology

Basic techniques such as electrophoresis, plasmid DNA isolation, PCR and RT-PCR reaction, molecular cloning etc. were performed according to (460) or manuals provided by the manufacturer.

5.2.2 Conditions of yeast cultures

Yeast strains were grown at 25°C or 30°C, either in YPD or YPGal medium (1% yeast extract, 2% Bacto Peptone, 2% glucose or 2% galactose, respectively) or in synthetic complete medium (0.77 g/l CSM medium, MP, 6.7 g/l yeast nitrogen base (YNB), Difco, 2% glucose, when necessary additionally supplemented with amino acids or nucleotide bases, BD Biosciences). For solid media agar was added to 2% concentration. Liquid cultures were grown to early logarithmic phase (OD_{600} 0.2–0.6) unless stated otherwise. Cells were harvested by centrifugation, washed with distilled water, frozen in liquid nitrogen and stored at -80°C. *GAL::CBP20*, *GAL::CBP80* and *GAL::eIF4E* strains were grown in l.3 l of complete synthetic medium w/o uracil with 2% sodium-L- lactate (0.77 g/l CSM –URA, MP, 6.7 g/l YNB w/o amino acids, Difco, 2% sodium-L- lactate, 50 μg/ml Ampicillin), at 30°C to OD_{600} 0.5–1. Induction with 650 mL of induction medium (60 g/l Bacto Peptone, Difco, 30 g/l Bacto Yeast Extract, Difco, 6% D-galactose, 50 mg/mL Ampicillin) was carried out for 6 h at 30°C.

5.2.3 Preparation of yeast strains

Myc- and HA-tagging was performed by one-step PCR procedure (461) using plasmids pFA6a-13Myc-natMX6 and pFA6a-3HA-natMX6 (459). BY4741 genomic DNA was used as a PCR template. Homologous sequences that determined the insertion region were added on primers overhangs (50 nt from the 3' ORF upstream the stop codon and 50 nt from the sequence just downstream the stop codon). Transformation procedure was as described (462) using 15-20 OD_{600} units of logarithmic phase yeast. Cells were harvested, washed with 10mM Tris-HCl pH 7.5 and incubated with buffer LiT (10mM Tris-HCl pH 7.5, 10 mM Lithium Acetate) supplemented with fresh DTT (10 mM) for 40 min. Cells were centrifuged and resuspended in transformation buffer (5-8 OD_{600} units per single transformation in 50 µl LiT, 20-100 µl purified PCR product, 5 µl of carrier DNA 10 µg/µl). After 10 min incubation in room temperature, 300 µl of 50% PEG 4000 (Merck) in LiT was added and incubation was continued for 10 min, followed by addition of 50 µl of DMSO and incubation for 15 min at 42°C. Cells were then centrifuged, resuspended in 1-20 ml of YPD medium and incubated at 25°C or 30°C for 1-12 h. After centrifugation, cells were plated on YPD Nat (100 mg/l, WERNER BioAgents). Cassette integration was confirmed using PCR and western blot analysis.

5.2.4 Isolation of yeast genomic DNA

2 ml of thick yeast culture were centrifuged and resuspended in buffer TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA) with 3% SDS and 500 µl glass beads (BioSpec). Cells were incubated for 10 min on vortex mixer. 400 µl TE and 500 µl phenol solution was added and the mixture was vortexed for another 5 min and centrifuged. Aqueous phase was twice extracted with cold phenol/chlorophorm/isoamyl alcohol mixture (25:24:1) and once with chlorophorm/isoamyl alcohol mixture (24:1). DNA was precipitated with 0.3 M ammonium acetate (pH 5.3) in 75% ethanol. After centrifugation and washing with 75% ethanol, DNA pellets were air-dried and resuspended in 50 µl of deionized water.

5.2.5 Isolation of RNA

Total RNA from yeast cells was isolated using a hot phenol procedure (463). Yeast cells (10-15 OD_{600} units) harvested at early logarithmic phase were resuspended in 400 buffer AE (50 mM sodium acetate, 10 mM EDTA) with 10% SDS and 400 µl of acidic phenol (pH 4-5). Suspension was incubated at 65°C with shaking for 5 min and then frozen in liquid nitrogen.

After centrifugation, aqueous phase was twice extracted with cold phenol/chlorophorm/isoamyl alcohol mixture (25:24:1) and once with chlorophorm/isoamyl alcohol mixture (24:1). RNA was precipitated with 0.3 M ammonium acetate (pH 5.3) in 75% ethanol. After centrifugation and washing with 75% ethanol, RNA pellets were air-dried and resuspended in 50 μl of deionized water.

5.2.6 Northern blot analysis

Northern hybridization was performed essentially as described (464). 8 μg of total RNA or 10 μg of RNA treated with 0.5 U of RNase H (Ambion) in the presence of oligonucleotide snR68son was mixed with the equal volume of loading buffer (98% formamide, 0.025% xylene cyanol, 0.025% bromophenol blue), incubated at 95°C for 5 min and separated on 6% denaturing polyacrylamide-urea gel in TBE buffer (89 mM Tris base, 89 mM boric acid, 2mM EDTA) at constant power 15 W. RNA was then transferred onto N+ nylon membrane (GE Healthcare) by electric transfer in 0.5 × TBE at 10 mA overnight. After fixing with UV crosslinking (1200 kJ/cm^2) blots was pre-hybridized with 5' end γ-^{32}P-labelled oligonucleotide probe (20 nt) in 20 ml of Denhardt buffer (5 × Denhardt reagent: 0.1% Ficoll 400, 0.1% polyvinylpyrrolidone, 0.1% BSA, 6 × SSPE: 60 mM phosphate buffer pH 7.4, 1 M NaCl, 6 mM EDTA, and 0.5% SDS) at 37°C for 30 min. Oligonucleotide was labelled using 1 U of PNK (NEB) and 20 μCi [γ-^{32}P]ATP (Hartman Analytics) in manufacturer buffer. 15 μl reaction was carried out at 37°C for 30 min. Hybridization was performed at 37°C for at least 8 h and the membrane was washed three times for 10 min at 42°C with 6 × SSPE buffer and exposed to Storage Phosphor Screen BAS-IP MS (Fuji Lifesciences). Autoradiographic images were obtained using laser scanner Typhoon FLA 9000 (GE Healthcare) at 650 nm wavelength and analysed using Image Quant (Molecular Dynamics).

5.2.7 RNase H treatment

20 μg of RNA was mixed with 1 pmol of specific oligonucleotide (snR68son for northern blot or cRTPCR1R for cRT-PCR analyses) in 10 μl of hybridisation buffer (25 mM Tris-HCl pH 7.5; 1 mM EDTA; 50 mM NaCl). RNA was denatured at 70°C for 10 min and slowly cooled down to 30°C. After that reaction buffer (final concentration: 20 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT, 30 mg/ml BSA) and 0.5 U of RNase H (Ambion) were added and the reaction was incubated at 30°C for 1 h. RNA was then extracted with phenol/chlorophorm/isoamyl alcohol mixture (25:24:1) and chlorophorm/isoamyl alcohol

mixture (24:1) and precipitated in 75% ethanol using 0.3 M ammonium acetate (pH 5.3) and 20 μg of glycogen.

5.2.8 cRT-PCR

cRT-PCR was carried out as described (251) with several modifications. 5' end of snR68 with the cap structure was removed prior to circularization either by RNase H (Ambion) treatment of 20 μg of total RNA in the presence of oligonucleotide cRTPCR1R (see 5.2.7) or by Tobacco Acid Pyrophosphatase (TAP, Epicentre) treatment for 2 h at 37°C using 8 μg of total RNA and 2.5 U of TAP followed by purification using RNeasy MinElute Cleanup Kit (Qiagen). Ligation reaction was incubated at 37°C for 90 min. After ligation RNA was purified either by extraction with phenol/chlorophorm/isoamyl alcohol mixture (25:24:1) and chlorophorm/isoamyl alcohol mixture (24:1) and subsequent precipitation or using RNeasy MinElute Cleanup Kit (Qiagen). Circular molecules were amplified by RT-PCR using Super Script Reverse transcriptase (Thermo Fisher Scientific) and 0.5-2 μg of RNA, 2 pmol of specific primer (RTsp) or 200 ng of random primers for the loading control. RNA mixed with primer and 1 mM dNTPs were denatured at 65°C for 5 min. After cooling on ice, the reaction mix was added to each sample (1 × manufacturer reaction buffer, 50 mM DTT, 20 U Ribolock (Thermo Fisher Scientific) and 100 U of Super Script Reverse transcriptase (Thermo Fisher Scientific)). Reaction was incubated at 50°C for 1 h. cDNA obtained using RTsp primer was diluted 4 times and the loading control 10 times prior to subsequent semi-quantitative PCR analysis. Products were cloned into pJET1.2 vector (Thermo Fisher Scientific) and sequenced at the DNA Sequencing and Oligonucleotides Synthesis facility (Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw). Preparation of plasmid DNA was carried out by MSc Katarzyna Kowalska (IBB PAS, Warsaw) using 96 Well Plate Plasmid DNA Mini-Preps Kit and Janus JANUS Automated Workstation (PerkinElmer).

5.2.9 Expression and purification of recombinant proteins

Mouse eukaryotic initiation factor eIF4E (residues 28–217) was expressed in *Escherichia coli* BL21(DE3) as inclusion bodies by Dr Joanna Żuberek (Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw). The guanidinium-solubilized protein was renatured by one-step dialysis and purified by ion-exchange chromatography on a HiTrap SP column without contact with cap analogues by Dr Joanna Zuberek as described in (465). Concentration of eIF4E was determined

spectrophotometrically assuming ε_{280} = 53,400 M^{-1} cm^{-1}. Human DcpS was expressed in *E. coli* Rosetta DE3 according to the procedures described previously (466) with several modifications by Dr Zbigniew Darżynkiewicz (Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw). His-tagged DcpS was purified by MSc Krystian Stoduś (IBB PAS, Warsaw) by a two-step procedure on Ni-NTA Superflow Cartridge (Qiagen) followed by gel filtration on Superdex 200 column (GE Healthcare) using the AKTA Purifier system (GE Healthcare). Proteins were eluted with buffer IPP150 (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% rTX100). Concentration of DcpS was determined spectrophotometrically assuming ε_{280} = 30,495 M^{-1} cm^{-1}.

Nrd1/Nab3 complex ($Nrd1^{(1-548)}$ with N-terminal His_9 tag and $Nab3^{(191-556)}$ with N-terminal MBP tag) was purified as described (274) with several modifications. The Nrd1–Nab3 heterodimer was expressed in *E. coli* strain BL21 grown in ZYM-5052 autoinduction medium (10 g/l N-Z amine AS, 5 g/l yeast extract, 1 mM $MgSO_4$, 1 × trace elements mix[1], 1 × 5052[2], 1 × M[3], 100 μg/l kanamycin, 34 mg/l chloramphenicol) at 4°C for 3 days. The culture was harvested and resuspended in buffer L1 (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol, 10% glycerol, 0.1% Triton X100) and sonicated in a Bioruptor sonicator (Diagenode) using 35–30 s pulses at the high level. The lysate was then filtered using syringe filters (Sarstedt). Fusion proteins were purified by a two-step procedure on Ni-NTA Superflow Cartridge (Qiagen) followed by ion exchange chromatography on HiTrap Heparin HP (GE Healthcare) using the AKTA Purifier system (GE Healthcare). After protein loading, Ni-NTA Superflow resin was washed with buffer W1 (20 mM Tris-HCl pH 8.0, 1 M NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol). Proteins were eluted with buffer E1 gradient (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 600 mM imidazole, 10 mM β-mercaptoethanol, 100% in 40 min) and then dialysed into buffer D2 (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT, 1 mM PMSF, 20% glycerol). After sample loading, heparin resin was washed with buffer W1 (10 mM Tris-HCl pH 8.0, 150 mM NaCl) with 1 mM DTT. Proteins were eluted with buffer E2 gradient (10 mM Tris-HCl pH 8.0, 1.2 M NaCl, 1 mM DTT, 100% in 40 min). Alternatively, a second batch of the Nrd1/Nab3 complex was purified by MSc Krystian Stoduś (IBB PAS, Warsaw) using Ni-NTA chromatography followed by gel filtration on Superdex 200 column (GE Healthcare). Proteins were eluted with buffer IPP150.

[1] 1000 × trace elements mix contains: 50 mM $FeCl_3$ × 6 H_2O; 20 mM $CaCl_2$ × 2 H_2O; 10 mM $MnCl_2$ × 4 H_2O; 10 mM $ZnSO_4$ × 7 H_20; 2 mM $CoCl_2$ × $6H_2O$; 2 mM $CuCl_2$ × 2 H_2O; 2 mM $NiCl_2$ × 6 H_2O; 2 mM Na_2MoO_4 × 2 H_2O; 2 mM H_3BO_3.

[2] 50 × 5052 contains: 25% glycerol; 2.5% glucose; 10% lactose.

[3] 50 × M contains: 2.5 M NH_4Cl; 1.25 M KH_2PO_4; 1.25 M Na_2HPO_4 × 2 H_2O; 0.25 M Na_2SO_4.

Affinity tags were not removed. Concentration of the purified complex was estimated using Bradford method assuming 400 – 1000 μg ml^{-1}.

Since Cbp80 expression in bacterial system was not feasible, CBC complex was purified from the yeast Y258 strain expressing Cbp20-His_6/HA/ZZ and Cbp80-FLAG-2×Strep-tagII under the control of the P_{GAL1} promoter from a BG1805 or p425GAL1 plasmids, respectively (450, 458). Cloning and tagging of Cbp80 was performed by MSc Dorota Adamska during the course of her MSc project. After induction (see 5.2.2) yeast were harvested, resuspended in an equal volume of buffer L2 (40 mM Hepes pH 8.0, 150 mM NaCl, 1 mM β-mercaptoethanol, 0.2% Tween20, protease inhibitors (Roche), 1 mM PMSF), frozen in liquid nitrogen and homogenized in laboratory blender (Waring) with dry ice (4 × 2 min). The lysate was centrifuged at 50000×g for 20 min and at 120000×g for 1h 30 min at 4°C. The extract was dialyzed for 4 h at 4°C in buffer D1 (40 mM Hepes pH 8.0, 150 mM NaCl, 1 mM β-mercaptoethanol, 1 mM PMSF, 20% glycerol). CBC was isolated on IgG Sepharose (GE Healthcare), after extensive washes with buffers IPP500 (10 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.1% Triton X100) and 3C (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM Sodium Citrate pH 8.0, 0.1% Triton X100, 10 mM β-mercaptoethanol), proteins were eluted overnight by protease 3C treatment. The eluate was applied on Superdex 200 column (GE Healthcare) and further purified using the AKTA Purifier system (GE Healthcare). Proteins were eluted with buffer IPP150. Concentration of the purified complex was estimated using Bradford method assuming 350 μg ml^{-1}.

5.2.10 Preparation of whole-cell extracts

Yeast extract for evaluation of modified cap affinity resins was prepared from 2 l of yeast culture (*CBP20-Myc* and *GAL::eIF4E*) at OD_{600} 1-2. Cell pellets were resuspended in an equal volume of buffer L3 (40 mM Hepes pH 8.0, 250 mM NaCl, 1 mM DTT), frozen in liquid nitrogen and homogenized in laboratory blender (Waring) with dry ice (4 × 2 min). The lysate was then centrifuged for 20 min at 50000×g at 4°C. The extract was cleared by ultracentrifugation at 120000×g for 1h 30 min at 4°C and dialyzed for 4 h at 4°C in buffer D3 (40 mM Hepes pH 8.0, 150 mM NaCl, 1 mM DTT, 1 mM PMSF, 20% glycerol).

Extracts used for protein binding studies were prepared from 250 ml of yeast culture (*CBP20-Myc*, *CBP80-Myc* and *RAD30-Myc*) at OD_{600} 0.6-0.8. Cell pellet was frozen in liquid nitrogen, resuspended in 1 ml of buffer L4 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.2% Tween 20, 1 mM β-mercaptoethanol, protease inhibitors (Roche), 1 mM PMSF) and lysed with 500 μl of glass beads (BioSpec). The lysate was centrifuged at 3200×g for 5 min at 4°C

and the supernatant was further clarified by a 30 min spin at 16000×g at 4°C. The extract was dialysed for 2 h at 4°C in buffer D4 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.2% Tween 20, 1 mM β-mercaptoethanol, 1 mM PMSF, 20% glycerol).

5.2.11 Purification of cap-binding proteins using cap modified affinity resins

Yeast extract (containing Cbp20-Myc or eIF4E-His6/HA/ZZ) or protein samples (520 µg/ml eIF4E and 200 µg/ml DcpS) were supplemented with reduced Triton X100 (0.1% final concentration) and GTP (100 µM final concentration). 300–500 µl of extract or protein sample were incubated for 2 h with 50 µl (settled volume) of resin (**1**–**4**) equilibrated with buffer IPP150 in 4°C. Flow through fractions were separated and resins were washed twice with 0. Bi; 4 × 40 µl) by incubation with the resin for 5 min. The purification procedure was carried out in 0.8 ml spin columns (Mo Bi Tec). 8 µl of collected fractions was separated by SDS-PAGE and in the case of purification from the extract were then analysed by western blot using anti-His and anti-Myc antibodies (see 5.2.13).

5.2.12 Protein interaction assays

Pull-down assays were carried out using whole cell yeast extracts prepared from *CBP20-Myc*, *CBP80-Myc* and *RAD30-Myc* strains and the recombinant Nrd1/Nab3 complex. The mixture of *CBP20-Myc* (30 µl) and *CBP80-Myc* (100 µl) extracts or *RAD30-Myc (*150 µl) extract were incubated with or without (negative control) approximately 60 µg of Nrd1/Nab3 proteins in buffer B (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM β-mercaptoethanol) on a rotating wheel for 1 h at 4°C. For RNase treatment, the reaction was incubated with or without 1 µg of RNase A (Sigma) for 2 h at 4°C prior to adding Nrd1/Nab3 proteins. RNase treatment was carried out by MSc Dorota Adamska during the course of her master project. The mixture was transferred to spin columns (Mo Bi Tec) and bound for 1 h at 4°C to 25 µl of Ni-NTA-Agarose (Qiagen) pre-blocked with 30 µg of BSA. Flow-through fractions were separated by centrifugation (1 min at 400×g), the resin was washed three times with 700 µl of buffer W2 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 20 mM imidazole, 1 mM β-mercaptoethanol) and bound proteins were eluted with buffer E4 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 250 mM imidazole, 1 mM β-mercaptoethanol) for 15 min at room temperature. 8 µl of collected fractions was separated by SDS-PAGE and analysed by western blot using anti-His and anti-Myc antibodies (see 5.2.13).

For *in vitro* binding, approximately 20 µg of purified CBC was incubated with 25 µg of recombinant Nrd1/Nab3 in 200 µl of IPP150 buffer on a rotating wheel at 4°C for 1 h. The

mixture was incubated at 4°C for 1 h in spin columns (Mo Bi Tec) with 1 mM GTP (Thermo Scientific) and 50 µl of m^7GTP-Sepharose equilibrated with IPP150. Bound proteins were purified as described for purification of cap binding proteins (see 5.2.11). 10 µl of collected fractions was analysed on 10% SDS-PAGE gel and stained with SYPRO Ruby Protein Gel Stain (Invitrogen).

5.2.13 Western blot analysis

After SDS-PAGE electrophoresis, proteins were transferred at nitrocellulose membrane (GE Healthcare) using semi-dry blotter (Bio-Rad) in buffer Tris-glycine (25 mM Tris-HCl, 250 mM glycine) with 20% methanol at 15 V for 1 h. Proteins were stained with Ponceau S dye in 3% acetic acid (the dye was then removed with distilled water). Membranes were blocked in 5% milk in PBST buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH2HPO4, 0.05% Tween 20) for 1 h at room temperature or overnight at 4°C. The blots were incubated with appropriate primary antibody: mouse anti-Myc antibody (Santa Cruz Biotechnology, 1:500 dilution), mouse anti-His antibody (Calbiochem, 1:3000 dilution), mouse anti-HA antibody (Covance, 1:1000 dilution) or PAP, peroxidase-anti-peroxidase antibody (1:5000). After three washes with PBST with 5% milk, membrane was incubated with secondary anti-mouse horseradish peroxidase-conjugated antibody (Calbiochem, 1:10000 dilution). After two washes in PBST with 5% milk and one with PBST, visualization was performed using Enhanced Chemiluminescence protocol (ECL) using ECL reagent (100 mM Tris-HCl pH 8.0, 225 µM coumaric acid, 1.25 mM luminol, 0.015% H_20_2) and CCD-camera (Fluorchem SP, Gel Biosciences).

5.2.14 Chromatin immunoprecipitation (ChIP)

ChIP assay was based on (251) with several modifications. 50 ml of yeast culture (OD_{600} 0.4-0.5) was crosslinked with 1% (v/v) formaldehyde at room temperature for 20 min and quenched with 375 mM glycine. Cells (20 OD_{600} units) were resuspended in 1 ml of cold FA1-lysis buffer (50 mM HEPES-KOH pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, protease inhibitors (Roche) and 2 mM PMSF) and disrupted with 300 µl of glass beads (BioSpec) using Fast Prep apparatus (MP Biomedicals), two times for 40 s at 6 m/s with 5 min rest period between runs. After glass beads separation, the lysate was diluted with 1 ml of FA1-lysis buffer and sonicated in a Bioruptor sonicator (Diagenode) using 35–30 s pulses at the high level. The lysate was clarified by 40 min spin at 16000×g at 4°C. 500 µl of extract diluted 5 times with FA-1 buffer was incubated overnight at 4°C with

20 µl of Dynabeads Protein G (Invitrogen) and appropriate amount of antibodies (10 µl anti-Myc (Santa Cruz Biotechnology), 4 µl anti-HA (Covance), 1 µl anti-Rbp3 (Neoclone), 5 µl anti-Ser7-P clone 24E12, 20 µl anti-Ser5-P clone 3E8 and 20 µl anti-Ser2-P clone 3E10 (CTD antibodies were a kind gift from Prof. Dirk Eick). 20 µl of the input sample were frozen in -80°C. Beads were washed twice at room temperature with 500 µl of FA1-lysis buffer, FA2-lysis buffer (50 mM HEPES-KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, and 0.1% sodium deoxycholate), ChIP wash buffer (10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 0.5% Nonidet P-40, 0.5% sodium deoxycholate) and TE. All washes were performed at room temperature. Beads were resuspended in 100 µl of ChIP elution buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 1% SDS) and samples, including 20 µl of the input sample, were incubated with 40 µg of Proteinase K (Bioline) for 2 h at 56°C and 5 h at 65°C to digest proteins, reverse cross-linking and elute DNA. For RNAse treatment, diluted extract was incubated with or without RNase A (10 U, Qiagen) and T1 (500 U, Roche) at 37°C for 1 h. Extracts were then incubated for 2 h at room temperature with 10 µl of Dynabeads and 5 µl of anti-Myc antibody (Santa Cruz Biotechnology). Washes were performed as described above. Beads resuspended in elution buffer with Proteinase K were incubated 2 h at 56°C and 12 h at 65°C. DNA was purified using the commercial clean-up kit (Axygen). Additional wash step was applied to completely remove traces of SDS. DNA was eluted twice with 35 µl of elution buffer. Prior to qPCR analysis DNA samples, except no-antibody control, were diluted 3 to 10 times. Reaction mixes for some experiments were pipetted using JANUS Integrator automated workstation (PerkinElmer). Real time PCR analysis was performed using Roche Light Cycler 480 with 4 to 6 primers pairs for each gene. Primers efficiencies were assayed in each run and were higher than 80%. Immunoprecipitation efficiency was expressed as percent of the input, calculated using DNA concentrations as established by the Light Cycler software (20). Data represent the results of at least three independent biological replicates. The background level of DNA non-specifically bound to magnetic beads was negligible as determined by immunoprecipitations without antibodies or using anti-Myc and anti-HA antibodies with the extract from wild-type (BY4741) cells.

6. LITERATURE

1. Costa,F.F. (2010) Non-coding RNAs: Meet thy masters. *BioEssays*, **32**, 599–608.
2. Egloff,S., Dienstbier,M. and Murphy,S. (2012) Updating the RNA polymerase CTD code: Adding gene-specific layers. *Trends Genet.*, **28**, 333–41.
3. Hsin,J. P. and Manley,J.L. (2012) The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.*, **26**, 2119–37.
4. Heidemann,M., Hintermair,C., Voß,K. and Eick,D. (2013) Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim. Biophys. Acta*, **1829**, 55–62.
5. Napolitano,G., Lania,L. and Majello,B. (2014) RNA polymerase II CTD modifications: how many tales from a single tail. *J. Cell. Physiol.*, **229**, 538–44.
6. Lenasi,T. and Barboric,M. (2013) Mutual relationships between transcription and pre-mRNA processing in the synthesis of mRNA. *Wiley Interdiscip. Rev. RNA*, **4**, 139–54.
7. Meinhart,A., Kamenski,T., Hoeppner,S., Baumli,S. and Cramer,P. (2005) A structural perspective of CTD function. *Genes Dev.*, **19**, 1401–15.
8. Kim,H., Erickson,B., Luo,W., Seward,D., Graber,J.H., Pollock,D.D., Megee,P.C. and Bentley,D.L. (2010) Gene-specific RNA polymerase II phosphorylation and the CTD code. *Nat. Struct. Mol. Biol.*, **17**, 1279–86.
9. Tietjen,J.R., Zhang,D.W., Rodríguez-Molina,J.B., White,B.E., Akhtar,M.S., Heidemann,M., Li,X., Chapman,R.D., Shokat,K., Keles,S., *et al.* (2010) Chemical-genomic dissection of the CTD code. *Nat. Struct. Mol. Biol.*, **17**, 1154–61.
10. Mayer,A., Lidschreiber,M., Siebert,M., Leike,K., Söding,J. and Cramer,P. (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nat. Struct. Mol. Biol.*, **17**, 1272–78.
11. Nojima,T., Gomes,T., Grosso,A.R.F., Kimura,H., Dye,M.J., Dhir,S., Carmo-Fonseca,M. and Proudfoot,N.J. (2015) Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell*, **161**, 526–40.
12. Galbraith,M.D., Donner,A.J. and Espinosa,J.M. (2010) CDK8: a positive regulator of transcription. *Transcription*, **1**, 4–12.
13. Mosley,A.L., Pattenden,S.G., Carey,M., Venkatesh,S., Gilmore,J.M., Florens,L., Workman,J.L. and Washburn,M.P. (2009) Rtr1 Is a CTD phosphatase that regulates RNA polymerase II during the transition from Serine 5 to Serine 2 phosphorylation. *Mol. Cell*, **34**, 168–78.
14. Zhang,D.W., Mosley,A.L., Ramisetty,S.R., Rodríguez-Molina,J.B., Washburn,M.P. and Ansari,A.Z. (2012) Ssu72 phosphatase-dependent erasure of phospho-Ser7 marks on the RNA polymerase II C-terminal domain is essential for viability and transcription termination. *J. Biol. Chem.*, **287**, 8541–51.
15. Bataille,A.R., Jeronimo,C., Jacques,P.É., Laramée,L., Fortin,M.È., Forest,A., Bergeron,M., Hanes,S.D. and Robert,F. (2012) A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. *Mol. Cell*, **45**, 158–70.
16. Ghazy,M.A, He,X., Singh,B.N., Hampsey,M. and Moore,C. (2009) The essential N terminus of the Pta1 scaffold protein is required for snoRNA transcription termination and Ssu72 function but is dispensable for pre-mRNA 3'-end processing. *Mol. Cell. Biol.*, **29**, 2296–307.
17. Schwer,B. and Shuman,S. (2011) Deciphering the RNA polymerase II CTD code in fission yeast. *Mol. Cell*, **43**, 311–8.
18. Kim,M., Suh,H., Cho,E.J. and Buratowski,S. (2009) Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2,5, and 7. *J. Biol. Chem.*, **284**, 26421–426.
19. Egloff,S., Zaborowska,J., Laitem,C., Kiss,T. and Murphy,S. (2012) Ser7 phosphorylation of the CTD recruits the RPAP2 Ser5 phosphatase to snRNA genes. *Mol. Cell*, **45**, 111–22.
20. Egloff,S., O'Reilly,D., Chapman,R.D., Taylor,A., Tanzhaus,K., Pitts,L., Eick,D. and Murphy,S. (2007) Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene

expression. *Science*, **318**, 1777–9.

21. Egloff,S., Al-Rawaf,H., O'Reilly,D. and Murphy,S. (2009) Chromatin structure is implicated in 'late' elongation checkpoints on the U2 snRNA and beta-actin genes. *Mol. Cell. Biol.*, **29**, 4002–13.
22. Egloff,S., Szczepaniak,S.A., Dienstbier,M., Taylor,A., Knight,S. and Murphy,S. (2010) The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J. Biol. Chem.*, **285**, 20564–69.
23. Buratowski,S. (2009) Progression through the RNA Polymerase II CTD Cycle. *Mol. Cell*, **36**, 541–6.
24. Qiu,H., Hu,C. and Hinnebusch,A.G. (2009) Phosphorylation of the Pol II CTD by KIN28 enhances BUR1/BUR2 recruitment and Ser2 CTD phosphorylation near promoters. *Mol. Cell*, **33**, 752–62.
25. Zhou,K., Kuo,W.H.W., Fillingham,J. and Greenblatt,J.F. (2009) Control of transcriptional elongation and cotranscriptional histone modification by the yeast BUR kinase substrate Spt5. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 6956–61.
26. Liu,Y., Warfield,L., Zhang,C., Luo,J., Allen,J., Lang,W.H., Ranish,J., Shokat,K.M. and Hahn,S. (2009) Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Mol. Cell. Biol.*, **29**, 4852–63.
27. Brès,V., Yoh,S.M. and Jones,K.A. (2008) The multi-tasking P-TEFb complex. *Curr. Opin. Cell Biol.*, **20**, 334–40.
28. Bartkowiak,B., Liu,P., Phatnani,H.P., Fuda,N.J., Cooper,J.J., Price,D.H., Adelman,K., Lis,J.T. and Greenleaf,A.L. (2010) CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.*, **24**, 2303–16.
29. Bartkowiak,B. and Greenleaf,A.L. (2011) Phosphorylation of RNAPII: To P-TEFb or not to P-TEFb? *Transcription*, **2**, 115–19.
30. Devaiah,B.N., Lewis,B.A., Cherman,N., Hewitt,M.C., Albrecht,B.K., Robey,P.G., Ozato,K., Sims,R.J. and Singer,D.S. (2012) BRD4 is an atypical kinase that phosphorylates serine2 of the RNA polymerase II carboxy-terminal domain. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 6927–32.
31. Lenasi,T., Peterlin,B.M. and Barboric,M. (2011) Cap-binding protein complex links pre-mRNA capping to transcription elongation and alternative splicing through positive transcription elongation factor b (P-TEFb). *J. Biol. Chem.*, **286**, 22758–68.
32. Hossain,M.A., Chung,C., Pradhan,S.K. and Johnson,T.L. (2013) The yeast cap binding complex modulates transcription factor recruitment and establishes proper histone H3K36 trimethylation during active transcription. *Mol. Cell. Biol.*, **33**, 785–99.
33. Lidschreiber,M., Leike,K. and Cramer,P. (2013) Cap completion and C-terminal repeat domain kinase recruitment underlie the initiation-elongation transition of RNA polymerase II. *Mol. Cell. Biol.*, **33**, 3805–16.
34. Clemente-Blanco,A., Sen,N., Mayan-Santos,M., Sacristán,M.P., Graham,B., Jarmuz,A., Giess,A., Webb,E., Game,L., Eick,D., *et al.* (2011) Cdc14 phosphatase promotes segregation of telomeres through repression of RNA polymerase II transcription. *Nat. Cell Biol.*, **13**, 1450–6.
35. Lunde,B.M., Reichow,S.L., Kim,M., Suh,H., Leeper,T.C., Yang,F., Mutschler,H., Buratowski,S., Meinhart,A. and Varani,G. (2010) Cooperative interaction of transcription termination factors with the RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.*, **17**, 1195–201.
36. Ursic,D., Chinchilla,K., Finkel,J.S. and Culbertson,M.R. (2004) Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing. *Nucleic Acids Res.*, **32**, 2441–52.
37. Chinchilla,K., Rodriguez-Molina,J.B., Ursic,D., Finkel,J.S., Ansari,A.Z. and Culbertsona,M.R. (2012) Interactions of Sen1, Nrd1, and Nab3 with multiple phosphorylated forms of the Rpb1 C-terminal domain in Saccharomyces cerevisiae. *Eukaryot. Cell*, **11**, 417–29.
38. Jimeno-Gonzalez,S., Schmid,M., Malagon,F., Haaning,L.L. and Jensen,T.H. (2014) Rat1p maintains RNA polymerase II CTD phosphorylation balance. *RNA*, **20**, 551–8.
39. Eick,D. and Geyer,M. (2013) The RNA polymerase II carboxy-terminal domain (CTD) code.

Chem. Rev., **113**, 8456–90.

40. Lenstra,T.L., Tudek,A., Clauder,S., Xu,Z., Pachis,S.T., Van Leenen,D., Kemmeren,P., Steinmetz,L.M., Libri,D. and Holstege,F.C.P. (2013) The role of Ctk1 kinase in termination of small non-coding RNAs. *PLoS One*, **8**, e80495.
41. Grzechnik,P., Gdula,M.R. and Proudfoot,N.J. (2015) Pcf11 orchestrates transcription termination pathways in yeast. *Genes Dev.*, **29**, 849–61.
42. Mayer,A., Heidemann,M., Lidschreiber,M., Schreieck,A., Sun,M., Hintermair,C., Kremmer,E., Eick,D. and Cramer,P. (2012) CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. *Science*, **336**, 1723–5.
43. Hintermair,C., Heidemann,M., Koch,F., Descostes,N., Gut,M., Gut,I., Fenouil,R., Ferrier,P., Flatley,A., Kremmer,E., *et al.* (2012) Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. *EMBO J.*, **31**, 2784–97.
44. Hsin,J.P., Sheth,A. and Manley,J.L. (2011) RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. *Science*, **334**, 683–6.
45. Baskaran,R., Dahmus,M.E. and Wang,J.Y. (1993) Tyrosine phosphorylation of mammalian RNA polymerase II carboxyl-terminal domain. *Proc. Natl. Acad. Sci. U. S. A.*, **90**, 11167–71.
46. Schreieck,A., Easter,A.D., Etzold,S., Wiederhold,K., Lidschreiber,M., Cramer,P. and Passmore,L.A. (2014) RNA polymerase II termination involves C-terminal-domain tyrosine dephosphorylation by CPF subunit Glc7. *Nat. Struct. Mol. Biol.*, **21**, 175–9.
47. Hsu,P.L., Yang,F., Smith-Kinnaman,W., Yang,W., Song,J.E., Mosley,A.L. and Varani,G. (2014) Rtr1 is a dual specificity phosphatase that dephosphorylates Tyr1 and Ser5 on the RNA polymerase II CTD. *J. Mol. Biol.*, **426**, 2970–81.
48. Noble,C.G., Hollingworth,D., Martin,S.R., Ennis-Adeniran,V., Smerdon,S.J., Kelly,G., Taylor,I.A. and Ramos,A. (2005) Key features of the interaction between Pcf11 CID and RNA polymerase II CTD. *Nat. Struct. Mol. Biol.*, **12**, 144–51.
49. Werner-Allen,J.W., Lee,C.J., Liu,P., Nicely,N.I., Wang,S., Greenleaf,A.L. and Zhou,P. (2011) cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. *J. Biol. Chem.*, **286**, 5717–26.
50. Kubicek,K., Cerna,H., Holub,P., Dev,G., Pasulka,J., Hrossova,D., Loehr,F., Hofr,C. and Vanacova,S. (2012) Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. *Genes Dev.*, **26**, 1891–6.
51. Singh,N., Ma,Z., Gemmill,T., Wu,X., DeFiglio,H., Rossettini,A., Rabeler,C., Beane,O., Morse,R.H., Palumbo,M.J., *et al.* (2009) The Ess1 prolyl isomerase is required for transcription termination of small noncoding RNAs via the Nrd1 pathway. *Mol. Cell*, **36**, 255–66.
52. Xiang,K., Nagaike,T., Xiang,S., Kilic,T., Beh,M.M., Manley,J.L. and Tong,L. (2010) Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. *Nature*, **467**, 729–33.
53. Gemmill,T.R., Wu,X. and Hanes,S.D. (2005) Vanishingly low levels of Ess1 prolyl-isomerase activity are sufficient for growth in Saccharomyces cerevisiae. *J. Biol. Chem.*, **280**, 15510–7.
54. Sims,R.J., Rojas,L.A., Beck,D., Bonasio,R., Schüller,R., Drury,W.J., Eick,D. and Reinberg,D. (2011) The C-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science*, **332**, 99–103.
55. Coudreuse,D., van Bakel,H., Dewez,M., Soutourina,J., Parnell,T., Vandenhaute,J., Cairns,B., Werner,M. and Hermand,D. (2010) A gene-specific requirement of RNA polymerase II CTD phosphorylation for sexual differentiation in S. pombe. *Curr. Biol.*, **20**, 1053–64.
56. Viladevall,L., St Amour,C. V, Rosebrock,A., Schneider,S., Zhang,C., Allen,J.J., Shokat,K.M., Schwer,B., Leatherwood,J.K. and Fisher,R.P. (2009) TFIIH and P-TEFb coordinate transcription with capping enzyme recruitment at specific genes in fission yeast. *Mol. Cell*, **33**, 738–51.
57. West,M.L. and Corden,J.L. (1995) Construction and analysis of yeast RNA polymerase II CTD deletion and substitution mutations. *Genetics*, **140**, 1223–33.
58. McCracken,S., Fong,N., Yankulov,K., Ballantyne,S., Pan,G., Greenblatt,J., Patterson,S.D., Wickens,M. and Bentley,D.L. (1997) The C-terminal domain of RNA polymerase II couples

mRNA processing to transcription. *Nature*, **385**, 357–61.

59. McCracken,S., Fong,N., Rosonina,E., Yankulov,K., Brothers,G., Siderovski,D., Hessel,A., Foster,S., Shuman,S. and Bentley,D.L. (1997) 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.*, **11**, 3306–18.
60. Stiller,J.W., McConaughy,B.L. and Hall,B.D. (2000) Evolutionary complementation for polymerase II CTD function. *Yeast*, **16**, 57–64.
61. Chapman,R.D., Heidemann,M., Albert,T.K., Mailhammer,R., Flatley,A., Meisterernst,M., Kremmer,E. and Eick,D. (2007) Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science*, **318**, 1780–2.
62. Stiller,J.W. and Cook,M.S. (2004) Functional unit of the RNA polymerase II C-terminal domain lies within heptapeptide pairs. *Eukaryot. Cell*, **3**, 735–40.
63. Liu,P., Kenney,J.M., Stiller,J.W. and Greenleaf,A.L. (2010) Genetic organization, length conservation, and evolution of RNA polymerase II carboxyl-terminal domain. *Mol. Biol. Evol.*, **27**, 2628–41.
64. Perales,R. and Bentley,D. (2009) 'Cotranscriptionality': the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell*, **36**, 178–91.
65. Rando,O.J. and Winston,F. (2012) Chromatin and transcription in yeast. *Genetics*, **190**, 351–87.
66. Tous,C., Rondón,A.G., García-Rubio,M., González-Aguilera,C., Luna,R. and Aguilera,A. (2011) A novel assay identifies transcript elongation roles for the Nup84 complex and RNA processing factors. *EMBO J.*, **30**, 1953–64.
67. Tomson,B.N. and Arndt,K.M. (2013) The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. *Biochim. Biophys. Acta*, **1829**, 116–26.
68. Sheldon,K.E., Mauger,D.M. and Arndt,K.M. (2005) A requirement for the Saccharomyces cerevisiae Paf1 complex in snoRNA 3' end formation. *Mol. Cell*, **20**, 225–36.
69. Tomson,B.N., Crisucci,E.M., Heisler,L.E., Gebbia,M., Nislow,C. and Arndt,K.M. (2013) Effects of the Paf1 complex and histone modifications on snoRNA 3'-end formation reveal broad and locus-specific regulation. *Mol. Cell. Biol.*, **33**, 170–82.
70. Terzi,N., Churchman,L.S., Vasiljeva,L., Weissman,J. and Buratowski,S. (2011) H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Mol. Cell. Biol.*, **31**, 3569–83.
71. Alén,C., Kent,N.A., Jones,H.S., O'Sullivan,J., Aranda,A. and Proudfoot,N.J. (2002) A role for chromatin remodeling in transcriptional termination by RNA polymerase II. *Mol. Cell*, **10**, 1441–52.
72. Morillon,A., Karabetsou,N., O'Sullivan,J., Kent,N., Proudfoot,N. and Mellor,J. (2003) Isw1 chromatin remodeling ATPase coordinates transcription elongation and termination by RNA polymerase II. *Cell*, **115**, 425–435.
73. Topisirovic,I., Svitkin,Y. V., Sonenberg,N. and Shatkin,A.J. (2011) Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip. Rev. RNA*, **2**, 277–298.
74. Gonatopoulos- Pournatzis,T. and Cowling,V.H. (2014) Cap-binding complex (CBC). *Biochem. J.*, **457**, 231–42.
75. Suh,M.H., Meyer,P.A., Gu,M., Ye,P., Zhang,M., Kaplan,C.D., Lima,C.D. and Fu,J. (2010) A dual interface determines the recognition of RNA polymerase II by RNA capping enzyme. *J. Biol. Chem.*, **285**, 34027–38.
76. Ghosh,A., Shuman,S. and Lima,C.D. (2011) Structural insights to how mammalian capping enzyme reads the CTD code. *Mol. Cell*, **43**, 299–310.
77. Lahudkar,S., Durairaj,G., Uprety,B. and Bhaumik,S.R. (2013) A novel role for Cet1p mRNA 5'-triphosphatase in promoter proximal accumulation of RNA polymerase II in saccharomyces cerevisiase. *Genetics*, **196**, 161–76.
78. Myers,L.C., Lacomis,L., Erdjument-Bromage,H. and Tempst,P. (2002) The yeast capping enzyme

represses RNA polymerase II transcription. *Mol. Cell*, **10**, 883–94.
79. Jimeno-González,S., Haaning,L.L., Malagon,F. and Jensen,T.H. (2010) The yeast 5′-3′ exonuclease Rat1p functions during transcription elongation by RNA polymerase II. *Mol. Cell*, **37**, 580–7.
80. Jiao,X., Xiang,S., Oh,C., Martin,C.E., Tong,L. and Kiledjian,M. (2010) Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature*, **467**, 608–11.
81. Izaurralde,E., Lewis,J., McGuigan,C., Jankowska,M., Darzynkiewicz,E. and Mattaj,I.W. (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell*, **78**, 657–68.
82. Calero,G., Wilson,K.F., Ly,T., Rios-Steiner,J.L., Clardy,J.C. and Cerione,R.A. (2002) Structural basis of m7GpppG binding to the nuclear cap-binding protein complex. *Nat. Struct. Biol.*, **9**, 912–7.
83. Mazza,C., Segref,A., Mattaj,I.W. and Cusack,S. (2002) Large-scale induced fit recognition of an m7GpppG cap analogue by the human nuclear cap-binding complex. *EMBO J.*, **21**, 5548–57.
84. Fortes,P., Kufel,J., Fornerod,M., Polycarpou-Schwarz,M., Lafontaine,D., Tollervey,D. and Mattaj,I.W. (1999) Genetic and physical interactions involving the yeast nuclear cap-binding complex. *Mol. Cell. Biol.*, **19**, 6543–53.
85. Narita,T., Yung,T.M.C., Yamamoto,J., Tsuboi,Y., Tanabe,H., Tanaka,K., Yamaguchi,Y. and Handa,H. (2007) NELF interacts with CBC and participates in 3′ end processing of replication-dependent histone mRNAs. *Mol. Cell*, **26**, 349–65.
86. Wong,C.M., Qiu,H., Hu,C., Dong,J. and Hinnebusch,A.G. (2007) Yeast cap binding complex impedes recruitment of cleavage factor IA to weak termination sites. *Mol. Cell. Biol.*, **27**, 6520–31.
87. Pabis,M., Neufeld,N., Steiner,M.C., Bojic,T., Shav-Tal,Y. and Neugebauer,K.M. (2013) The nuclear cap-binding complex interacts with the U4/U6·U5 tri-snRNP and promotes spliceosome assembly in mammalian cells. *RNA*, **19**, 1054–63.
88. Visa,N., Izaurralde,E., Ferreira,J., Daneholt,B. and Mattaj,I.W. (1996) A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. *J. Cell Biol.*, **133**, 5–14.
89. Lewis,J.D., Izaurralde,E., Jarmolowski,A., McGuigan,C. and Mattaj,I.W. (1996) A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.*, **10**, 1683–98.
90. Lewis,J.D. and Izaurralde,E. (1997) The role of the cap structure in RNA processing and nuclear export. *Eur. J. Biochem.*, **247**, 461–9.
91. Colot,H. V., Stutz,F. and Rosbash,M. (1996) The yeast splicing factor Mud13p is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. *Genes Dev.*, **10**, 1699–708.
92. Görnemann,J., Kotovic,K.M., Hujer,K. and Neugebauer,K.M. (2005) Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Mol. Cell*, **19**, 53–63.
93. Zhang,D. and Rosbash,M. (1999) Identification of eight proteins that cross-link to pre-mRNA in the yeast commitment complex. *Genes Dev.*, **13**, 581–92.
94. Flaherty,S.M., Fortes,P., Izaurralde,E., Mattaj,I.W. and Gilmartin,G.M. (1997) Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc. Natl. Acad. Sci. U. S. A.*, **94**, 11893–8.
95. Andersen,P.R., Domanski,M., Kristiansen,M.S., Storvall,H., Ntini,E., Verheggen,C., Schein,A., Bunkenborg,J., Poser,I., Hallais,M., *et al.* (2013) The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat. Struct. Mol. Biol.*, **20**, 1367–76.
96. Hallais,M., Pontvianne,F., Andersen,P.R., Clerici,M., Lener,D., Benbahouche,N.E.H., Gostan,T., Vandermoere,F., Robert,M.C., Cusack,S., *et al.* (2013) CBC-ARS2 stimulates 3'-end maturation of multiple RNA families and favors cap-proximal processing. *Nat. Struct. Mol. Biol.*, **20**, 1358–66.
97. Lahudkar,S., Shukla,A., Bajwa,P., Durairaj,G., Stanojevic,N. and Bhaumik,S.R. (2011) The

mRNA cap-binding complex stimulates the formation of pre-initiation complex at the promoter via its interaction with Mot1p in vivo. *Nucleic Acids Res.*, **39**, 2188–209.

98. Isken,O. and Maquat,L.E. (2007) Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. *Genes Dev.*, **21**, 1833–56.
99. Das,B., Guo,Z., Russo,P., Chartrand,P. and Sherman,F. (2000) The role of nuclear cap binding protein Cbc1p of yeast in mRNA termination and degradation. *Mol. Cell. Biol.*, **20**, 2827–38.
100. Das,B., Butler,J.S. and Sherman,F. (2003) Degradation of normal mRNA in the nucleus of Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **23**, 5502–15.
101. Das,S., Saha,U. and Das,B. (2014) Cbc2p, Upf3p and eIF4G are components of the DRN (Degradation of mRNA in the Nucleus) in Saccharomyces cerevisiae. *FEMS Yeast Res.*, **14**, 922–32.
102. Tseng,C.K., Wang,H.F., Burns,A.M., Schroeder,M.R., Gaspari,M. and Baumann,P. (2015) Human telomerase RNA processing and quality control. *Cell Rep.*, doi: 10.1016/j.celrep.2015.10.075.
103. Vasiljeva,L. and Buratowski,S. (2006) Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. *Mol. Cell*, **21**, 239–48.
104. Dias,S.M.G., Wilson,K.F., Rojas,K.S., Ambrosio,A.L.B. and Cerione,R.A. (2009) The molecular basis for the regulation of the cap-binding complex by the importins. *Nat. Struct. Mol. Biol.*, **16**, 930–7.
105. Schwer,B., Chang,J. and Shuman,S. (2013) Structure-function analysis of the 5' end of yeast U1 snRNA highlights genetic interactions with the Msl5*Mud2 branchpoint-binding complex and other spliceosome assembly factors. *Nucleic Acids Res.*, **41**, 7485–500.
106. Qiu,Z.R., Chico,L., Chang,J., Shuman,S. and Schwer,B. (2012) Genetic interactions of hypomorphic mutations in the m7G cap-binding pocket of yeast nuclear cap binding complex: An essential role for Cbc2 in meiosis via splicing of MER3 pre-mRNA. *RNA*, **18**, 1996–2011.
107. Schwer,B., Erdjument-Bromage,H. and Shuman,S. (2011) Composition of yeast snRNPs and snoRNPs in the absence of trimethylguanosine caps reveals nuclear cap binding protein as a gained U1 component implicated in the cold-sensitivity of tgs1Δ cells. *Nucleic Acids Res.*, **39**, 6715–28.
108. Mouaikel,J., Verheggen,C., Bertrand,E., Tazi,J. and Bordonné,R. (2002) Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell*, **9**, 891–901.
109. Matera, a G., Terns,R.M. and Terns,M.P. (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat. Rev. Mol. Cell Biol.*, **8**, 209–20.
110. Franke,J., Gehlen,J. and Ehrenhofer-Murray,A.E. (2008) Hypermethylation of yeast telomerase RNA by the snRNA and snoRNA methyltransferase Tgs1. *J. Cell Sci.*, **121**, 3553–60.
111. Hausmann,S. and Shuman,S. (2005) Specificity and mechanism of RNA cap guanine-N2 methyltransferase (Tgs1). *J. Biol. Chem.*, **280**, 4021–4.
112. Chang,J., Schwer,B. and Shuman,S. (2010) Mutational analyses of trimethylguanosine synthase (Tgs1) and Mud2: proteins implicated in pre-mRNA splicing. *RNA*, **16**, 1018–31.
113. Matera,A. G. and Shpargel,K.B. (2006) Pumping RNA: nuclear bodybuilding along the RNP pipeline. *Curr. Opin. Cell Biol.*, **18**, 317–34.
114. Mouaikel,J., Bujnicki,J.M., Tazi,J. and Bordonné,R. (2003) Sequence-structure-function relationships of Tgs1, the yeast snRNA/snoRNA cap hypermethylase. *Nucleic Acids Res.*, **31**, 4899–909.
115. Worch,R., Niedzwiecka,A., Stepinski,J., Mazza,C., Jankowska-Anyszka,M., Darzynkiewicz,E., Cusack,S. and Stolarski,R. (2005) Specificity of recognition of mRNA 5' cap by human nuclear cap-binding complex. *RNA*, **11**, 1355–63.
116. Hausmann,S., Zheng,S., Costanzo,M., Brost,R.L., Garcin,D., Boone,C., Shuman,S. and Schwer,B. (2008) Genetic and biochemical analysis of yeast and human cap trimethylguanosine synthase: Functional overlap of 2,2,7-trimethylguanosine caps, small nuclear ribonucleoprotein components, PRE-mRNA splicing factors, and RNA decay pathways. *J. Biol. Chem.*, **283**,

31706–18.
117. Qiu,Z.R., Shuman,S. and Schwer,B. (2011) An essential role for trimethylguanosine RNA caps in Saccharomyces cerevisiae meiosis and their requirement for splicing of SAE3 and PCH2 meiotic pre-mRNAs. *Nucleic Acids Res.*, **39**, 5633–46.
118. Colau,G., Thiry,M., Leduc,V., Bordonné,R. and Lafontaine,D.L.J. (2004) The small nucle(ol)ar RNA cap trimethyltransferase is required for ribosome synthesis and intact nucleolar morphology. *Mol. Cell. Biol.*, **24**, 7976–86.
119. Chanfreau,G., Legrain,P. and Jacquier, a (1998) Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J. Mol. Biol.*, **284**, 975–88.
120. Lee,C.Y., Lee,A. and Chanfreau,G. (2003) The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of S. cerevisiae box C/D snoRNAs. *RNA*, **9**, 1362–70.
121. MacRae,I.J. and Doudna,J.A. (2007) Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr. Opin. Struct. Biol.*, **17**, 138–45.
122. Lamontagne,B., Tremblay, a and Abou Elela,S. (2000) The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. *Mol. Cell. Biol.*, **20**, 1104–15.
123. Liang,Y.H., Lavoie,M., Comeau,M.A., Abou Elela,S. and Ji,X. (2014) Structure of a eukaryotic RNase III postcleavage complex reveals a double-ruler mechanism for substrate selection. *Mol. Cell*, **54**, 431–44.
124. Chanfreau,G., Buckle,M. and Jacquier, a (2000) Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 3142–47.
125. Nagel,R. and Ares,M. (2000) Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5'-AGNN-3' tetraloop. *RNA*, **6**, 1142–56.
126. Ghazal,G., Ge,D., Gervais-Bird,J., Gagnon,J. and Abou Elela,S. (2005) Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. *Mol. Cell. Biol.*, **25**, 2981–94.
127. Gaudin,C., Ghazal,G., Yoshizawa,S., Abou Elela,S. and Fourmy,D. (2006) Structure of an AAGU tetraloop and its contribution to substrate selection by yeast RNase III. *J. Mol. Biol.*, **363**, 322–31.
128. Wang,Z., Hartman,E., Roy,K., Chanfreau,G. and Feigon,J. (2011) Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. *Structure*, **19**, 999–1010.
129. Abou Elela,S., Igel,H. and Ares,M. (1996) RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell*, **85**, 115–24.
130. Kufel,J., Dichtl,B. and Tollervey,D. (1999) Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3' ETS but not the 5' ETS. *RNA*, **5**, 909–17.
131. Chanfreau,G., Rotondo,G., Legrain,P. and Jacquier,A. (1998) Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *EMBO J.*, **17**, 3726–37.
132. Henras,A.K., Bertrand,E. and Chanfreau,G. (2004) A cotranscriptional model for 3'-end processing of the Saccharomyces cerevisiae pre-ribosomal RNA precursor. *RNA*, **10**, 1572–85.
133. Catala,M., Lamontagne,B., Larose,S., Ghazal,G. and Abou Elela,S. (2004) Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. *Mol. Biol. Cell*, **15**, 3015–30.
134. Giorgi,C., Fatica,A., Nagel,R. and Bozzoni,I. (2001) Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease. *EMBO J.*, **20**, 6856–65.
135. Tremblay,A., Lamontagne,B., Catala,M., Yam,Y., Larose,S., Good,L. and Abou Elela,S. (2002) A physical interaction between Gar1p and Rnt1p Is required for the nuclear import of H/ACA small nucleolar RNA-associated proteins. *Mol. Cell. Biol.*, **22**, 4792–802.
136. Danin-Kreiselman,M., Lee,C.Y. and Chanfreau,G. (2003) RNAse III-mediated degradation of

unspliced pre-mRNAs and lariat introns. *Mol. Cell*, **11**, 1279–89.

137. Egecioglu,D.E., Kawashima,T.R. and Chanfreau,G.F. (2012) Quality control of MATa1 splicing and exon skipping by nuclear RNA degradation. *Nucleic Acids Res.*, **40**, 1787–96.
138. Larose,S., Laterreur,N., Ghazal,G., Gagnon,J., Wellinger,R.J. and Abou Elela,S. (2007) RNase III-dependent regulation of yeast telomerase. *J. Biol. Chem.*, **282**, 4373–81.
139. Lee,A., Henras,A.K. and Chanfreau,G. (2005) Multiple RNA surveillance pathways limit aberrant expression of iron uptake mRNAs and prevent iron toxicity in S. cerevisiae. *Mol. Cell*, **19**, 39–51.
140. Catala,M., Aksouh,L. and Abou Elela,S. (2012) RNA-dependent regulation of the cell wall Stress response. *Nucleic Acids Res.*, **40**, 7507–17.
141. Ghazal,G., Gagnon,J., Jacques,P.É., Landry,J.R., Robert,F. and Abou Elela,S. (2009) Yeast RNase III triggers polyadenylation-independent transcription termination. *Mol. Cell*, **36**, 99–109.
142. Rondón,A.G., Mischo,H.E., Kawauchi,J. and Proudfoot,N.J. (2009) Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. *Mol. Cell*, **36**, 88–98.
143. Abou Elela,S. and Ares,M. (1998) Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA. *EMBO J.*, **17**, 3738–46.
144. Reeder,R.H., Guevara,P. and Roan,J.G. (1999) Saccharomyces cerevisiae RNA polymerase I terminates transcription at the Reb1 terminator in vivo. *Mol. Cell. Biol.*, **19**, 7369–76.
145. Prescott,E.M., Osheim,Y.N., Jones,H.S., Alen,C.M., Roan,J.G., Reeder,R.H., Beyer,A.L. and Proudfoot,N.J. (2004) Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 6068–73.
146. Catala,M., Tremblay,M., Samson,E., Conconi,A. and Abou Elela,S. (2008) Deletion of Rnt1p alters the proportion of open versus closed rRNA gene repeats in yeast. *Mol. Cell. Biol.*, **28**, 619–29.
147. Henry,Y., Wood,H., Morrissey,J.P., Petfalski,E., Kearsey,S. and Tollervey,D. (1994) The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO J.*, **13**, 2452–63.
148. Petfalski,E., Dandekar,T., Henry,Y. and Tollervey,D. (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.*, **18**, 1181–89.
149. Qu,L.H., Henras,A., Lu,Y.J., Zhou,H., Zhou,W.X., Zhu,Y.Q., Zhao,J., Henry,Y., Caizergues-Ferrer,M. and Bachellerie,J.P. (1999) Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.*, **19**, 1144–58.
150. Villa,T., Ceradini,F., Presutti,C. and Bozzoni,I. (1998) Processing of the intron-encoded U18 small nucleolar RNA in the yeast Saccharomyces cerevisiae relies on both exo- and endonucleolytic activities. *Mol. Cell. Biol.*, **18**, 3376–83.
151. Geerlings,T.H., Vos,J.C. and Raué,H.A. (2000) The final step in the formation of 25S rRNA in Saccharomyces cerevisiae is performed by 5'-3' exonucleases. *RNA*, **6**, 1698–703.
152. Amberg,D.C., Goldstein,A.L. and Cole,C.N. (1992) Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. *Genes Dev.*, **6**,1173-89.
153. Johnson,A.W. (1997) Rat1p and Xrn1p are functionally interchangeable exoribonucleases that are restricted to and required in the nucleus and cytoplasm, respectively. *Mol. Cell. Biol.*, **17**, 6122–30.
154. Xue,Y., Bai,X., Lee,I., Kallstrom,G., Ho,J., Brown,J., Stevens,A. and Johnson,A.W. (2000) Saccharomyces cerevisiae RAI1 (YGL246c) is homologous to human DOM3Z and encodes a protein that binds the nuclear exoribonuclease Rat1p. *Mol. Cell. Biol.*, **20**, 4006–15.
155. Di Segni,G., McConaughy,B.L., Shapiro,R.A., Aldrich,T.L. and Hall,B.D. (1993) TAP1, a yeast gene that activates the expression of a tRNA gene with a defective internal promoter. *Mol. Cell. Biol.*, **13**, 3424–33.
156. Kenna,M., Stevens,A., McCammon,M. and Douglas,M.G. (1993) An essential yeast gene with

homology to the exonuclease-encoding XRN1/KEM1 gene also encodes a protein with exoribonuclease activity. *Mol. Cell. Biol.*, **13**, 341–50.

157. Poole,T.L. and Stevens,A. (1995) Comparison of features of the RNase activity of 5'-exonuclease-1 and 5'-exonuclease-2 of Saccharomyces cerevisiae. *Nucleic Acids Symp. Ser.*, **33**, 79-81
158. Bernstein,J. and Toth,E.A. (2012) Yeast nuclear RNA processing. *World J. Biol. Chem.*, **3**, 7–26.
159. Wolin,S.L., Sim,S. and Chen,X. (2012) Nuclear noncoding RNA surveillance: Is the end in sight? *Trends Genet.*, **28**, 306–13.
160. Parker,R. (2012) RNA degradation in Saccharomyces cerevisae. *Genetics*, **191**, 671–702.
161. Xiang,S., Cooper-Morgan,A., Jiao,X., Kiledjian,M., Manley,J.L. and Tong,L. (2009) Structure and function of the 5'-3' exoribonuclease Rat1 and its activating partner Rai1. *Nature*, **458**, 784–8.
162. Kim,M., Krogan,N.J., Vasiljeva,L., Rando,O.J., Nedea,E., Greenblatt,J.F. and Buratowski,S. (2004) The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature*, **432**, 517–22.
163. El Hage,A., Koper,M., Kufel,J. and Tollervey,D. (2008) Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. *Genes Dev.*, **22**, 1069–81.
164. Luo,W., Johnson,A.W. and Bentley,D.L. (2006) The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: Implications for a unified allosteric-torpedo model. *Genes Dev.*, **20**, 954–65.
165. Luke,B., Panza,A., Redon,S., Iglesias,N., Li,Z. and Lingner,J. (2008) The Rat1p 5' to 3' exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae. *Mol. Cell*, **32**, 465–77.
166. Iglesias,N., Redon,S., Pfeiffer,V., Dees,M., Lingner,J. and Luke,B. (2011) Subtelomeric repetitive elements determine TERRA regulation by Rap1/Rif and Rap1/Sir complexes in yeast. *EMBO Rep.*, **12**, 587–93.
167. Toone,W.M., Johnson,A.L., Banks,G.R., Toyn,J.H., Stuart,D., Wittenberg,C. and Johnston,L.H. (1995) Rme1, a negative regulator of meiosis, is also a positive activator of G1 cyclin gene expression. *EMBO J.*, **14**, 5824–32.
168. Chen,C.Y.A. and Shyu,A. Bin (2011) Mechanisms of deadenylation-dependent decay. *Wiley Interdiscip. Rev. RNA*, **2**, 167–83.
169. Ling,S.H.M., Qamra,R. and Song,H. (2011) Structural and functional insights into eukaryotic mRNA decapping. *Wiley Interdiscip. Rev. RNA*, **2**, 193–208.
170. Arribas-Layton,M., Wu,D., Lykke-Andersen,J. and Song,H. (2013) Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochim. Biophys. Acta*, **1829**, 580–9.
171. van Dijk,E., Cougot,N., Meyer,S., Babajko,S., Wahle,E. and Séraphin,B. (2002) Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J.*, **21**, 6915–24.
172. Piccirillo,C., Khanna,R. and Kiledjian,M. (2003) Functional characterization of the mammalian mRNA decapping enzyme hDcp2. *RNA*, **9**, 1138–47.
173. Steiger,M., Carr-Schmid,A., Schwartz,D.C., Kiledjian,M. and Parker,R. (2003) Analysis of recombinant yeast decapping enzyme. *RNA*, **9**, 231–8.
174. She,M., Decker,C.J., Svergun,D.I., Round,A., Chen,N., Muhlrad,D., Parker,R. and Song,H. (2008) Structural basis of dcp2 recognition and activation by dcp1. *Mol. Cell*, **29**, 337–49.
175. She,M., Decker,C.J., Chen,N., Tumati,S., Parker,R. and Song,H. (2006) Crystal structure and functional analysis of Dcp2p from Schizosaccharomyces pombe. *Nat. Struct. Mol. Biol.*, **13**, 63–70.
176. Deshmukh,M. V, Jones,B.N., Quang-Dang,D.U., Flinders,J., Floor,S.N., Kim,C., Jemielity,J., Kalek,M., Darzynkiewicz,E. and Gross,J.D. (2008) mRNA decapping is promoted by an RNA-binding channel in Dcp2. *Mol. Cell*, **29**, 324–36.
177. Borja,M.S., Piotukh,K., Freund,C. and Gross,J.D. (2011) Dcp1 links coactivators of mRNA

decapping to Dcp2 by proline recognition. *RNA*, **17**, 278–90.

178. Fenger-Grøn,M., Fillman,C., Norrild,B. and Lykke-Andersen,J. (2005) Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Mol. Cell*, **20**, 905–15.
179. Xu,J., Yang,J.Y., Niu,Q.W. and Chua,N.H. (2006) Arabidopsis DCP2, DCP1, and VARICOSE form a decapping complex required for postembryonic development. *Plant Cell*, **18**, 3386–98.
180. Sheth,U. and Parker,R. (2003) Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science*, **300**, 805–8.
181. Franks,T.M. and Lykke-Andersen,J. (2008) The control of mRNA decapping and P-body formation. *Mol. Cell*, **32**, 605–15.
182. Hu,W., Sweet,T.J., Chamnongpol,S., Baker,K.E. and Coller,J. (2009) Co-translational mRNA decay in Saccharomyces cerevisiae. *Nature*, **461**, 225–9.
183. Hu,W., Petzold,C., Coller,J. and Baker,K.E. (2010) Nonsense-mediated mRNA decapping occurs on polyribosomes in Saccharomyces cerevisiae. *Nat. Struct. Mol. Biol.*, **17**, 244–7.
184. Kufel,J., Bousquet-Antonelli,C., Beggs,J.D. and Tollervey,D. (2004) Nuclear pre-mRNA decapping and 5' degradation in yeast require the Lsm2-8p complex. *Mol. Cell. Biol.*, **24**, 9646–57.
185. Grousl,T., Ivanov,P., Frýdlová,I., Vasicová,P., Janda,F., Vojtová,J., Malínská,K., Malcová,I., Nováková,L., Janosková,D., *et al.* (2009) Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, Saccharomyces cerevisiae. *J. Cell Sci.*, **122**, 2078–88.
186. Geisler,S., Lojek,L., Khalil,A.M., Baker,K.E. and Coller,J. (2012) Decapping of long noncoding RNAs regulates inducible genes. *Mol. Cell*, **45**, 279–91.
187. Cloutier,S.C., Wang,S., Ma,W.K., Petell,C.J. and Tran,E.J. (2013) Long noncoding RNAs promote transcriptional poising of inducible genes. *PLoS Biol.*, **11**, 32–4.
188. Cohen,L., Mikhli,C., Jiao,X. and Kiledjian,M. (2005) Dcp2 decaps m2,2,7GpppN-capped RNAs, and its activity is sequence and context dependent. *Mol. Cell. Biol.*, **25**, 8779–91.
189. Song,M.G., Bail,S. and Kiledjian,M. (2013) Multiple Nudix family proteins possess mRNA decapping activity. *RNA*, **19**, 390–9.
190. Shukla,S. and Parker,R. (2014) Quality control of assembly-defective U1 snRNAs by decapping and 5'-to-3' exonucleolytic digestion. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, E3277–86.
191. Haimovich,G., Medina,D.A., Causse,S.Z., Garber,M., Millán-Zambrano,G., Barkai,O., Chávez,S., Pérez-Ortín,J.E., Darzacq,X. and Choder,M. (2013) Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. *Cell*, **153**, 1000–11.
192. Brannan,K., Kim,H., Erickson,B., Glover-cutter,K., Kim,S., Fong,N., Kiemele,L., Hansen,K., Davis,R., Lykke-Andersen,J., *et al.* (2012) mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol. Cell*, **46**, 311–24.
193. Tomasevic,N. and Peculis,B. (1999) Identification of a U8 snoRNA-specific binding protein. *J. Biol. Chem.*, **274**, 35914–20.
194. Ghosh,T., Peterson,B., Tomasevic,N. and Peculis,B.A. (2004) Xenopus U8 snoRNA binding protein is a conserved nuclear decapping enzyme. *Mol. Cell*, **13**, 817–28.
195. Song,M.G., Li,Y. and Kiledjian,M. (2010) Multiple mRNA decapping enzymes in mammalian cells. *Mol. Cell*, **40**, 423–32.
196. Chang,J.H., Jiao,X., Chiba,K., Oh,C., Martin,C.E., Kiledjian,M. and Tong,L. (2012) Dxo1 is a new type of eukaryotic enzyme with both decapping and 5′-3′ exoribonuclease activity. *Nat. Struct. Mol. Biol.*, **19**, 1011–17.
197. Jiao,X., Chang,J.H., Kilic,T., Tong,L. and Kiledjian,M. (2013) A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol. Cell*, **50**, 104–15.
198. Liu,H., Rodgers,N.D., Jiao,X. and Kiledjian,M. (2002) The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J.*, **21**, 4699–708.

199. Bail,S. and Kiledjian,M. (2008) DcpS, a general modulator of cap-binding protein-dependent processes? *RNA Biol.*, **5**, 216–9.
200. Beelman,C.A., Stevens,A., Caponigro,G., LaGrandeur,T.E., Hatfield,L., Fortner,D.M. and Parker,R. (1996) An essential component of the decapping enzyme required for normal rates of mRNA turnover. *Nature*, **382**, 642–6.
201. Dunckley,T. and Parker,R. (1999) The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. *EMBO J.*, **18**, 5411–22.
202. Anderson,J.S. and Parker,R.P. (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J.*, **17**, 1497–506.
203. van Hoof,A., Staples,R.R., Baker,R.E. and Parker,R. (2000) Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. *Mol. Cell. Biol.*, **20**, 8230–43.
204. Richard,P. and Manley,J.L. (2009) Transcription termination by nuclear RNA polymerases. *Genes Dev.*, **23**, 1247–69.
205. Kuehner,J.N., Pearson,E.L. and Moore,C. (2011) Unravelling the means to an end: RNA polymerase II transcription termination. *Nat. Rev. Mol. Cell Biol.*, **12**, 283–94.
206. Mischo,H.E. and Proudfoot,N.J. (2013) Disengaging polymerase: Terminating RNA polymerase II transcription in budding yeast. *Biochim. Biophys. Acta*, **1829**, 174–85.
207. Porrua,O. and Libri,D. (2015) Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat. Rev. Mol. Cell Biol.*, **16**, 190–202.
208. West,S. and Proudfoot,N.J. (2009) Transcriptional termination enhances protein expression in human cells. *Mol. Cell*, **33**, 354–64.
209. Andersen,P.K., Jensen,T.H. and Lykke-Andersen,S. (2013) Making ends meet: Coordination between RNA 3′-end processing and transcription initiation. *Wiley Interdiscip. Rev. RNA*, **4**, 233–46.
210. Lykke-Andersen,S. and Jensen,T.H. (2007) Overlapping pathways dictate termination of RNA polymerase II transcription. *Biochimie*, **89**, 1177–82.
211. Nedea,E., He,X., Kim,M., Pootoolal,J., Zhong,G., Canadien,V., Hughes,T., Buratowski,S., Moore,C.L. and Greenblatt,J. (2003) Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends. *J. Biol. Chem.*, **278**, 33000–10.
212. Garas,M., Dichtl,B. and Keller,W. (2008) The role of the putative 3' end processing endonuclease Ysh1p in mRNA and snoRNA synthesis. *RNA*, **14**, 2671–84.
213. Connelly,S. and Manley,J.L. (1988) A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.*, **2**, 440–52.
214. Logan,J., Falck-Pedersen,E., Darnell,J.E. and Shenk,T. (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. *Proc. Natl. Acad. Sci. U. S. A.*, **84**, 8306–10.
215. Birse,C.E., Minvielle-Sebastia,L., Lee,B.A., Keller,W. and Proudfoot,N.J. (1998) Coupling termination of transcription to messenger RNA maturation in yeast. *Science*, **280**, 298–301.
216. Dichtl,B., Blank,D., Sadowski,M. and Hu,W. (2002) Yhh1p/Cft1p directly links poly (A) site recognition and RNA polymerase II transcription termination. **21**, 4125–35.
217. Steinmetz,E.J. and Brow,D.A. (2003) Ssu72 protein mediates both poly(A)-coupled and poly(A)-independent termination of RNA polymerase II transcription. *Mol. Cell. Biol.*, **23**, 6339–49.
218. Kaneko,S., Rozenblatt-Rosen,O., Meyerson,M. and Manley,J.L. (2007) The multifunctional protein p54nrb/PSF recruits the exonuclease XRN2 to facilitate pre-mRNA 3' processing and transcription termination. *Genes Dev.*, **21**, 1779–89.
219. West,S., Proudfoot,N.J. and Dye,M.J. (2008) Molecular dissection of mammalian RNA polymerase II transcriptional termination. *Mol. Cell*, **29**, 600–10.
220. Schaughency,P., Merran,J. and Corden,J.L. (2014) Genome-wide mapping of yeast RNA polymerase II termination. *PLoS Genet.*, **10**, e1004632.

221. Pearson,E.L. and Moore,C.L. (2013) Dismantling promoter-driven RNA polymerase II transcription complexes in vitro by the termination factor Rat1. *J. Biol. Chem.*, **288**, 19750–9.

222. Park,J., Kang,M. and Kim,M. (2015) Unraveling the mechanistic features of RNA polymerase II termination by the 5'-3' exoribonuclease Rat1. *Nucleic Acids Res.*, **43**, 2625–37.

223. Mariconti,L., Loll,B., Schlinkmann,K., Wengi,A., Meinhart,A. and Dichtl,B. (2010) Coupled RNA polymerase II transcription and 3' end formation with yeast whole-cell extracts. *RNA*, **16**, 2205–17.

224. Fong,N., Brannan,K., Erickson,B., Kim,H., Cortazar,M.A., Sheridan,R.M., Nguyen,T., Karp,S. and Bentley,D.L. (2015) Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. *Mol. Cell*, **60**, 256–67.

225. Zhang,H., Rigo,F. and Martinson,H.G. (2015) Poly(A) signal-dependent transcription termination occurs through a conformational change mechanism that does not require cleavage at the poly(A) site. *Mol. Cell*, **59**, 437–48.

226. Licatalosi,D.D., Geiger,G., Minet,M., Schroeder,S., Cilli,K., McNeil,J.B. and Bentley,D.L. (2002) Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol. Cell*, **9**, 1101–11.

227. Sadowski,M., Dichtl,B., Hübner,W. and Keller,W. (2003) Independent functions of yeast Pcf11p in pre-mRNA 3′ end processing and in transcription termination. *EMBO J.*, **22**, 2167–77.

228. Kim,M., Vasiljeva,L., Rando,O.J., Zhelkovsky,A., Moore,C. and Buratowski,S. (2006) Distinct pathways for snoRNA and mRNA termination. *Mol. Cell*, **24**, 723–34.

229. Zhang,Z., Fu,J. and Gilmour,D.S. (2005) CTD-dependent dismantling of the RNA polymerase II elongation complex by the pre-mRNA 3′-end processing factor, Pcf11. *Genes Dev.*, **19**, 1572–80.

230. Amrani,N., Minet,M., Wyers,F., Dufour,M.E., Aggerbeck,L.P. and Lacroute,F. (1997) PCF11 encodes a third protein component of yeast cleavage and polyadenylation factor I. *Mol. Cell. Biol.*, **17**, 1102–09.

231. West,S. and Proudfoot,N.J. (2008) Human Pcf11 enhances degradation of RNA polymerase II-associated nascent RNA and transcriptional termination. *Nucleic Acids Res.*, **36**, 905–14.

232. Pearson,E. and Moore,C. (2014) The evolutionarily conserved Pol II flap loop contributes to proper transcription termination on short yeast genes. *Cell Rep.*, **9**, 821–8.

233. Bucheli,M.E. and Buratowski,S. (2005) Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. *EMBO J.*, **24**, 2150–60.

234. Bucheli,M.E., He,X., Kaplan,C.D., Moore,C.L. and Buratowski,S. (2007) Polyadenylation site choice in yeast is affected by competition between Npl3 and polyadenylation factor CFI. *RNA*, **13**, 1756–64.

235. Dermody,J.L., Dreyfuss,J.M., Villén,J., Ogundipe,B., Gygi,S.P., Park,P.J., Ponticelli,A.S., Moore,C.L., Buratowski,S. and Bucheli,M.E. (2008) Unphosphorylated SR-like protein Npl3 stimulates RNA polymerase II elongation. *PLoS One*, **3**. e3273.

236. Xu,C. and Henry,M.F. (2004) Nuclear export of hnRNP Hrp1p and nuclear export of hnRNP Npl3p are linked and influenced by the methylation state of Npl3p. *Mol. Cell. Biol.*, **24**, 10742–56.

237. McBride,A.E., Cook,J.T., Stemmler,E.A., Rutledge,K.L., McGrath,K.A. and Rubens,J.A. (2005) Arginine methylation of yeast mRNA-binding protein Npl3 directly affects its function, nuclear export, and intranuclear protein interactions. *J. Biol. Chem.*, **280**, 30888–98.

238. Wong,C.M., Tang,H.M.V., Kong,K.Y.E., Wong,G.W.O., Qiu,H., Jin,D.Y. and Hinnebusch,A.G. (2010) Yeast arginine methyltransferase Hmt1p regulates transcription elongation and termination by methylating Npl3p. *Nucleic Acids Res.*, **38**, 2217–28.

239. Noble,C.G., Walker,P.A., Calder,L.J. and Taylor,I.A. (2004) Rna14 - Rna15 assembly mediates the RNA-binding capability of Saccharomyces cerevisiae cleavage factor IA. *Nucleic Acids Res.*, **32**, 3364–75.

240. Kaida,D., Berg,M.G., Younis,I., Kasim,M., Singh,L.N., Wan,L. and Dreyfuss,G. (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature*, **468**, 664–8.

241. Andersen,P.K., Lykke-Andersen,S. and Jensen,T.H. (2012) Promoter-proximal polyadenylation sites reduce transcription activity. *Genes Dev.*, **26**, 2169–79.
242. Almada,A.E., Wu,X., Kriz,A.J., Burge,C.B. and Sharp,P.A. (2013) Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature*, **499**, 360–3.
243. Wu,X. and Sharp,P.A. (2013) Divergent transcription: A driving force for new gene origination? *Cell*, **155**, 990–6.
244. Steinmetz,E.J., Ng,S.B.H., Cloute,J.P. and Brow,D.A. (2006) cis- and trans-Acting determinants of transcription termination by yeast RNA polymerase II. *Mol. Cell. Biol.*, **26**, 2688–96.
245. Runner,V.M., Podolny,V. and Buratowski,S. (2008) The Rpb4 subunit of RNA polymerase II contributes to cotranscriptional recruitment of 3' processing factors. *Mol. Cell. Biol.*, **28**, 1883–91.
246. Kawauchi,J., Mischo,H., Braglia,P., Rondon,A. and Proudfoot,N.J. (2008) Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes Dev.*, **22**, 1082–92.
247. Steinmetz,E.J., Warren,C.L., Kuehner,J.N., Panbehi,B., Ansari,A.Z. and Brow,D.A. (2006) Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol. Cell*, **24**, 735–46.
248. Arndt,K.M. and Reines,D. (2014) Termination of transcription of short noncoding RNAs by RNA polymerase II. *Annu. Rev. Biochem.*, **84**, 381–404.
249. Fatica,A., Morlando,M. and Bozzoni,I. (2000) Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3'-processing apparatus. *EMBO J.*, **19**, 6218–29.
250. Morlando,M., Greco,P., Dichtl,B., Fatica,A., Keller,W. and Bozzoni,I. (2002) Functional analysis of yeast snoRNA and snRNA 3'-end formation mediated by uncoupling of cleavage and polyadenylation. *Mol. Cell. Biol.*, **22**, 1379–89.
251. Grzechnik,P. and Kufel,J. (2008) Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. *Mol. Cell*, **32**, 247–58.
252. Porrua,O., Hobor,F., Boulay,J., Kubicek,K., D'Aubenton-Carafa,Y., Gudipati,R.K., Stefl,R. and Libri,D. (2012) In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *EMBO J.*, **31**, 3935–48.
253. Miguel,A., Montón,F., Li,T., Gómez-Herreros,F., Chávez,S., Alepuz,P. and Pérez-Ortín,J.E. (2013) External conditions inversely change the RNA polymerase II elongation rate and density in yeast. *Biochim. Biophys. Acta*, **1829**, 1248–55.
254. Lemay,J.F., D'Amours,A., Lemieux,C., Lackner,D.H., St-Sauveur,V.G., Bähler,J. and Bachand,F. (2010) The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. *Mol. Cell*, **37**, 34–45.
255. Garland,W., Feigenbutz,M., Turner,M. and Mitchell,P. (2013) Rrp47 functions in RNA surveillance and stable RNA processing when divorced from the exoribonuclease and exosome-binding domains of Rrp6. *RNA*, **19**, 1659–68.
256. Schulz,D., Schwalb,B., Kiesel,A., Baejen,C., Torkler,P., Gagneur,J., Soeding,J. and Cramer,P. (2013) Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell*, **155**, 1075–87.
257. Wlotzka,W., Kudla,G., Granneman,S. and Tollervey,D. (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.*, **30**, 1790–803.
258. Webb,S., Hector,R.D., Kudla,G. and Granneman,S. (2014) PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. *Genome Biol.*, **15**, R8.
259. Arigo,J.T., Carroll,K.L., Ames,J.M. and Corden,J.L. (2006) Regulation of yeast NRD1 expression by premature transcription termination. *Mol. Cell*, **21**, 641–51.
260. Gudipati,R.K., Neil,H., Feuerbach,F., Malabat,C. and Jacquier,A. (2012) The yeast RPL9B gene is regulated by modulation between two modes of transcription termination. *EMBO J.*, **31**, 2427–37.

261. Wyers,F., Rougemaille,M., Badis,G., Rousselle,J.C., Dufour,M.E., Boulay,J., Régnault,B., Devaux,F., Namane,A., Séraphin,B., *et al.* (2005) Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell*, **121**, 725–37.

262. Davis,C.A. and Ares,M. (2006) Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 3262–7.

263. Castelnuovo,M., Zaugg,J.B., Guffanti,E., Maffioletti,A., Camblong,J., Xu,Z., Clauder-Münster,S., Steinmetz,L.M., Luscombe,N.M. and Stutz,F. (2014) Role of histone modifications and early termination in pervasive transcription and antisense-mediated gene silencing in yeast. *Nucleic Acids Res.*, **42**, 4348–62.

264. Peart,N., Sataluri,A., Baillat,D. and Wagner,E.J. (2013) Non-mRNA 3' end formation: How the other half lives. *Wiley Interdiscip. Rev. RNA*, **4**, 491–506.

265. Egloff,S., O'Reilly,D. and Murphy,S. (2008) Expression of human snRNA genes from beginning to end. *Biochem. Soc. Trans.*, **36**, 590–4.

266. O'Reilly,D., Kuznetsova,O. V., Laitem,C., Zaborowska,J., Dienstbier,M. and Murphy,S. (2014) Human snRNA genes use polyadenylation factors to promote efficient transcription termination. *Nucleic Acids Res.*, **42**, 264–75.

267. Yamamoto,J., Hagiwara,Y., Chiba,K., Isobe,T., Narita,T., Handa,H. and Yamaguchi,Y. (2014) DSIF and NELF interact with Integrator to specify the correct post-transcriptional fate of snRNA genes. *Nat. Commun.*, **5**, 4263.

268. Steinmetz,E.J. and Brow,D.A. (1996) Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1. *Mol. Cell. Biol.*, **16**, 6993–7003.

269. Steinmetz,E.J., Conrad,N.K., Brow,D.A. and Corden,J.L. (2001) RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. *Nature*, **413**, 327–31.

270. Conrad,N.K., Wilson,S.M., Steinmetz,E.J., Patturajan,M., Brow,D.A., Swanson,M.S. and Corden,J.L. (2000) A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. *Genetics*, **154**, 557–71.

271. Gudipati,R.K., Villa,T., Boulay,J. and Libri,D. (2008) Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat. Struct. Mol. Biol.*, **15**, 786–794.

272. Vasiljeva,L., Kim,M., Mutschler,H., Buratowski,S. and Meinhart,A. (2008) The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat. Struct. Mol. Biol.*, **15**, 795–804.

273. Heo,D.H., Yoo,I., Kong,J., Lidschreiber,M., Mayer,A., Choi,B.Y., Hahn,Y., Cramer,P., Buratowski,S. and Kim,M. (2013) The RNA polymerase II C-terminal domain-interacting domain of yeast Nrd1 contributes to the choice of termination pathway and couples to RNA processing by the nuclear exosome. *J. Biol. Chem.*, **288**, 36676–90.

274. Carroll,K.L., Ghirlando,R., Ames,J.M. and Corden,J.L. (2007) Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. *RNA*, **13**, 361–73.

275. Carroll,K.L., Pradhan,D.A., Granek,J.A., Clarke,N.D. and Corden,J.L. (2004) Identication of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. *Mol. Cell. Biol.*, **24**, 6241–52.

276. Thiebaut,M., Kisseleva-Romanova,E., Rougemaille,M., Boulay,J. and Libri,D. (2006) Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the Nrd1-Nab3 pathway in genome surveillance. *Mol. Cell*, **23**, 853–64.

277. Kuehner,J.N. and Brow,D.A. (2008) Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol. Cell*, **31**, 201–11.

278. Creamer,T.J., Darby,M.M., Jamonnak,N., Schaughency,P., Hao,H., Wheelan,S.J. and Corden,J.L. (2011) Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1. *PLoS Genet.*, **7**, e1002329.

279. Jamonnak,N., Creamer,T.J., Darby,M.M., Schaughency,P., Wheelan,S.J. and Corden,J.L. (2011)

Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA*, **17**, 2011–25.

280. Tudek,A., Porrua,O., Kabzinski,T., Lidschreiber,M., Kubicek,K., Fortova,A., Lacroute,F., Vanacova,S., Cramer,P., Stefl,R., *et al.* (2014) Molecular basis for coordinating transcription termination with noncoding RNA degradation. *Mol. Cell*, **55**, 467–81.
281. Chi,A., Huttenhower,C., Geer,L.Y., Coon,J.J., Syka,J.E.P., Bai,D.L., Shabanowitz,J., Burke,D.J., Troyanskaya,O.G. and Hunt,D.F. (2007) Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 2193–8.
282. Ficarro,S.B., McCleland,M.L., Stukenberg,P.T., Burke,D.J., Ross,M.M., Shabanowitz,J., Hunt,D.F. and White,F.M. (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. *Nat. Biotechnol.*, **20**, 301–5.
283. Darby,M.M., Serebreni,L., Pan,X., Boeke,J.D. and Corden,J.L. (2012) The Saccharomyces cerevisiae Nrd1-Nab3 transcription termination pathway acts in opposition to Ras signaling and mediates response to nutrient depletion. *Mol. Cell. Biol.*, **32**, 1762–75.
284. Wilson,S.M., Datar,K. V, Paddy,M.R., Swedlow,J.R. and Swanson,M.S. (1994) Characterization of nuclear polyadenylated RNA-binding proteins in Saccharomyces cerevisiae. *J. Cell Biol.*, **127**, 1173–84.
285. Loya,T.J., Rourke,T.W., Degtyareva,N. and Reines,D. (2013) A network of interdependent molecular interactions describes a higher order Nrd1-Nab3 complex involved in yeast transcription termination. *J. Biol. Chem.*, **288**, 34158–67.
286. O'Rourke,T.W., Loya,T.J., Head,P.E., Horton,J.R. and Reines,D. (2015) Amyloid-like assembly of the low complexity domain of yeast Nab3. *Prion*, **9**, 1–9.
287. Loya,T.J., O'Rourke,T.W. and Reines,D. (2012) A genetic screen for terminator function in yeast identifies a role for a new functional domain in termination factor Nab3. *Nucleic Acids Res.*, **40**, 7476–91.
288. Loya,T.J., O'Rourke,T.W. and Reines,D. (2013) Yeast Nab3 protein contains a self-assembly domain found in human heterogeneous nuclear ribonucleoprotein-C (hnRNP-C) that is necessary for transcription termination. *J. Biol. Chem.*, **288**, 2111–17.
289. Lunde,B.M., Hörner,M. and Meinhart,A. (2011) Structural insights into cis element recognition of non-polyadenylated RNAs by the Nab3-RRM. *Nucleic Acids Res.*, **39**, 337–46.
290. DeMarini,D.J., Winey,M., Ursic,D., Webb,F. and Culbertson,M.R. (1992) SEN1, a positive effector of tRNA-splicing endonuclease in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **12**, 2154–64.
291. Kim,H.D., Choe,J. and Seo,Y.S. (1999) The sen1(+) gene of Schizosaccharomyces pombe , a homologue of budding yeast SEN1, encodes an RNA and DNA helicase. *Biochemistry*, **38**, 14697–710.
292. Finkel,J.S., Chinchilla,K., Ursic,D. and Culbertson,M.R. (2010) Sen1p performs two genetically separable functions in transcription and processing of U5 small nuclear RNA in Saccharomyces cerevisiae. *Genetics*, **184**, 107–18.
293. Ghaemmaghami,S., Huh,W.K., Bower,K., Howson,R.W., Belle,A., Dephoure,N., O'Shea,E.K. and Weissman,J.S. (2003) Global analysis of protein expression in yeast. *Nature*, **425**, 737–41.
294. Yuryev,A., Patturajan,M., Litingtung,Y., Joshi,R. V, Gentile,C., Gebara,M. and Corden,J.L. (1996) The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **93**, 6975–80.
295. Ursic,D., Himmel,K.L., Gurley,K.A, Webb,F. and Culbertson,M.R. (1997) The yeast SEN1 gene is required for the processing of diverse RNA classes. *Nucleic Acids Res.*, **25**, 4778–85.
296. Chen,X., Muller,U., Sundling,K.E. and Brow,D.A. (2014) Saccharomyces cerevisiae Sen1 as a model for the study of mutations in human Senataxin that elicit cerebellar ataxia. *Genetics*, **198**, 577–90.
297. Porrua,O. and Libri,D. (2013) A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nat. Struct. Mol. Biol.*, **20**, 884–91.

298. Nedea,E., Nalbant,D., Xia,D., Theoharis,N.T., Suter,B., Richardson,C.J., Tatchell,K., Kislinger,T., Greenblatt,J.F. and Nagy,P.L. (2008) The Glc7 phosphatase subunit of the cleavage and polyadenylation factor Is essential for transcription termination on snoRNA genes. *Mol. Cell*, **29**, 577–87.

299. He,X. and Moore,C. (2005) Regulation of yeast mRNA 3' end processing by phosphorylation. *Mol. Cell*, **19**, 619–29.

300. Suraweera,A., Becherel,O.J., Chen,P., Rundle,N., Woods,R., Nakamura,J., Gatei,M., Criscuolo,C., Filla,A., Chessa,L., *et al.* (2007) Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage. *J. Cell Biol.*, **177**, 969–79.

301. De Amicis,A., Piane,M., Ferrari,F., Fanciulli,M., Delia,D. and Chessa,L. (2011) Role of senataxin in DNA damage and telomeric stability. *DNA Repair (Amst).*, **10**, 199–209.

302. Becherel,O.J., Yeo,A.J., Stellati,A., Heng,E.Y.H., Luff,J., Suraweera,A.M., Woods,R., Fleming,J., Carrie,D., McKinney,K., *et al.* (2013) Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing. *PLoS Genet.*, **9**, e1003435.

303. Hazelbaker,D.Z., Marquardt,S., Wlotzka,W. and Buratowski,S. (2013) Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Mol. Cell*, **49**, 55–66.

304. Mischo,H.E., Gómez-González,B., Grzechnik,P., Rondón,A.G., Wei,W., Steinmetz,L., Aguilera,A. and Proudfoot,N.J. (2011) Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell*, **41**, 21–32.

305. Chan,Y.A., Aristizabal,M.J., Lu,P.Y.T., Luo,Z., Hamza,A., Kobor,M.S., Stirling,P.C. and Hieter,P. (2014) Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-chip. *PLoS Genet.*, **10**, e1004288.

306. Alzu,A., Bermejo,R., Begnis,M., Lucca,C., Piccini,D., Carotenuto,W., Saponaro,M., Brambati,A., Cocito,A., Foiani,M., *et al.* (2012) Senataxin associates with replication forks to protect fork integrity across RNA-polymerase-II-transcribed genes. *Cell*, **151**, 835–46.

307. Skourti-Stathaki,K., Proudfoot,N.J. and Gromak,N. (2011) Human Senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell*, **42**, 794–805.

308. Chen,Y.Z., Bennett,C.L., Huynh,H.M., Blair,I.P., Puls,I., Irobi,J., Dierick,I., Abel,A., Kennerson,M.L., Rabin,B.A., *et al.* (2004) DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am. J. Hum. Genet.*, **74**, 1128–35.

309. Moreira,M.C., Klur,S., Watanabe,M., Németh,A.H., Le Ber,I., Moniz,J.C., Tranchant,C., Aubourg,P., Tazir,M., Schöls,L., *et al.* (2004) Senataxin, the ortholog of a yeast RNA helicase, is mutant in ataxia-ocular apraxia 2. *Nat. Genet.*, **36**, 225–7.

310. Winey,M. and Culbertson,M.R. (1988) Mutations affecting the tRNA-splicing endonuclease activity of Saccharomyces cerevisiae. *Genetics*, **118**, 609–17.

311. Page,B.D. and Snyder,M. (1992) CIK1: a developmentally regulated spindle pole body-associated protein important for microtubule functions in Saccharomyces cerevisiae. *Genes Dev.*, **6**, 1414–29.

312. Ursic,D., DeMarini,D.J. and Culbertson,M.R. (1995) Inactivation of the yeast Sen1 protein affects the localization of nucleolar proteins. *Mol. Gen. Genet.*, **249**, 571–84.

313. Carneiro,T., Carvalho,C., Braga,J., Rino,J., Milligan,L., Tollervey,D. and Carmo-Fonseca,M. (2007) Depletion of the yeast nuclear exosome subunit Rrp6 results in accumulation of polyadenylated RNAs in a discrete domain within the nucleolus. *Mol. Cell. Biol.*, **27**, 4157–65.

314. Dheur,S., Vo le,T.A., Voisinet-Hakil,F., Minet,M., Schmitter,J.M., Lacroute,F., Wyers,F. and Minvielle-Sebastia,L. (2003) Pti1p and Ref2p found in association with the mRNA 3′ end formation complex direct snoRNA maturation. *EMBO J.*, **22**, 2831–40.

315. Ganem,C., Devaux,F., Torchet,C., Jacq,C., Quevillon-Cheruel,S., Labesse,G., Facca,C. and Faye,G. (2003) Ssu72 is a phosphatase essential for transcription termination of snoRNAs and specific mRNAs in yeast. *EMBO J.*, **22**, 1588–98.

316. Morlando,M., Ballarino,M., Greco,P., Caffarelli,E., Dichtl,B. and Bozzoni,I. (2004) Coupling between snoRNP assembly and 3' processing controls box C/D snoRNA biosynthesis in yeast. *EMBO J.*, **23**, 2392–401.
317. Krishnamurthy,S., He,X., Reyes-Reyes,M., Moore,C. and Hampsey,M. (2004) Ssu72 is an RNA polymerase II CTD phosphatase. *Mol. Cell*, **14**, 387–94.
318. Cheng,H., He,X. and Moore,C.L. (2004) The essential WD repeat protein Swd2 has dual functions in RNA polymerase II transcription termination and lysine 4 methylation of histone H3. *Mol. Cell. Biol.*, **24**, 2932–43.
319. Dichtl,B., Aasland,R. and Keller,W. (2004) Functions for S. cerevisiae Swd2p in 3' end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. *RNA*, **10**, 965–77.
320. Soares,L.M. and Buratowski,S. (2012) Yeast Swd2 is essential because of antagonism between Set1 histone methyltransferase complex and APT (associated with Pta1) termination factor. *J. Biol. Chem.*, **287**, 15219–31.
321. Ballarino,M., Morlando,M., Pagano,F., Fatica,A. and Bozzoni,I. (2005) The cotranscriptional assembly of snoRNPs controls the biosynthesis of H/ACA snoRNAs in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **25(13)**, 5396–403.
322. Fox,M.J., Gao,H., Smith-Kinnaman,W.R., Liu,Y. and Mosley,A.L. (2015) The exosome component Rrp6 Is required for RNA polymerase II termination at specific targets of the Nrd1-Nab3 pathway. *PLOS Genet.*, **10**, e1004999.
323. Gudipati,R.K., Xu,Z., Lebreton,A., Séraphin,B., Steinmetz,L.M., Jacquier,A. and Libri,D. (2012) Extensive degradation of RNA precursors by the exosome in wild-type cells. *Mol. Cell*, **48**, 409–21.
324. Castelnuovo,M., Rahman,S., Guffanti,E., Infantino,V., Stutz,F. and Zenklusen,D. (2013) Bimodal expression of PHO84 is modulated by early termination of antisense transcription. *Nat. Struct. Mol. Biol.*, **20**, 851–8.
325. Lemay,J.-F., Larochelle,M., Marguerat,S., Atkinson,S., Bähler,J. and Bachand,F. (2014) The RNA exosome promotes transcription termination of backtracked RNA polymerase II. *Nat. Struct. Mol. Biol.*, **21**, 919–26.
326. Allmang,C., Kufel,J., Chanfreau,G., Mitchell,P., Petfalski,E. and Tollervey,D. (1999) Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.*, **18**, 5399–410.
327. van Hoof,A., Lennertz,P. and Parker,R. (2000) Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol. Cell. Biol.*, **20**, 441–52.
328. Chanfreau,G., Abou Elela,S., Ares,M. and Guthrie,C. (1997) Alternative 3'-end processing of U5 snRNA by RNase III. *Genes Dev.*, **11**, 2741–51.
329. Seipelt,R.L., Zheng,B., Asuru,A. and Rymond,B.C. (1999) U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway. *Nucleic Acids Res.*, **27**, 587–95.
330. Lund,M.K., Kress,T.L. and Guthrie,C. (2008) Autoregulation of Npl3, a yeast SR protein, requires a novel downstream region and serine phosphorylation. *Mol. Cell. Biol.*, **28**, 3873–81.
331. Lykke-Andersen,S., Tomecki,R., Jensen,T.H. and Dziembowski,A. (2011) The eukaryotic RNA exosome: same scaffold but variable catalytic subunits. *RNA Biol.*, **8**, 61–6.
332. Januszyk,K. and Lima,C.D. (2014) The eukaryotic RNA exosome. *Curr. Opin. Struct. Biol.*, **24**, 132–40.
333. Schneider,C. and Tollervey,D. (2013) Threading the barrel of the RNA exosome. *Trends Biochem. Sci.*, **38**, 485–93.
334. Liu,Q., Greimann,J.C. and Lima,C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223–37.
335. Dziembowski,A., Lorentzen,E., Conti,E. and Séraphin,B. (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.*, **14**, 15–22.
336. Chekanova,J.A., Dutko,J.A., Mian,I.S. and Belostotsky,D.A. (2002) Arabidopsis thaliana

exosome subunit AtRrp4p is a hydrolytic 3’-5' exonuclease containing S1 and KH RNA-binding domains. *Nucleic Acids Res.*, **30**, 695–700.

337. Bonneau,F., Basquin,J., Ebert,J., Lorentzen,E. and Conti,E. (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell*, **139**, 547–59.
338. Allmang,C., Petfalski,E., Podtelejnikov,A., Mann,M., Tollervey,D. and Mitchell,P. (1999) The yeast exosome and human PM-Scl are related complexes of 3’ - 5' exonucleases. *Genes Dev.*, **13**, 2148–58.
339. Tomecki,R., Kristiansen,M.S., Lykke-Andersen,S., Chlebowski,A., Larsen,K.M., Szczesny,R.J., Drazkowska,K., Pastula,A., Andersen,J.S., Stepien,P.P., *et al.* (2010) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.*, **29**, 2342–57.
340. Lebreton,A., Tomecki,R., Dziembowski,A. and Séraphin,B. (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature*, **456**, 993–6.
341. Schaeffer,D., Tsanova,B., Barbas,A., Reis,F.P., Dastidar,E.G., Sanchez-Rotunno,M., Arraiano,C.M. and van Hoof,A. (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nat. Struct. Mol. Biol.*, **16**, 56–62.
342. Schneider,C., Leung,E., Brown,J. and Tollervey,D. (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res.*, **37**, 1127–40.
343. Graham,A.C., Davis,S.M. and Andrulis,E.D. (2009) Interdependent nucleocytoplasmic trafficking and interactions of Dis3 with Rrp6, the core exosome and importin-alpha3. *Traffic*, **10**, 499–513.
344. Mamolen,M., Smith,A. and Andrulis,E.D. (2010) Drosophila melanogaster Dis3 N-terminal domains are required for ribonuclease activities, nuclear localization and exosome interactions. *Nucleic Acids Res.*, **38**, 5507–17.
345. Makino,D.L., Baumgärtner,M. and Conti,E. (2013) Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature*, **495**, 70–5.
346. Cristodero,M., Böttcher,B., Diepholz,M., Scheffzek,K. and Clayton,C. (2008) The Leishmania tarentolae exosome: purification and structural analysis by electron microscopy. *Mol. Biochem. Parasitol.*, **159**, 24–9.
347. Malet,H., Topf,M., Clare,D.K., Ebert,J., Bonneau,F., Basquin,J., Drazkowska,K., Tomecki,R., Dziembowski,A., Conti,E., *et al.* (2010) RNA channelling by the eukaryotic exosome. *EMBO Rep.*, **11**, 936–42.
348. Schneider,C., Kudla,G., Wlotzka,W., Tuck,A. and Tollervey,D. (2012) Transcriptome-wide analysis of exosome targets. *Mol. Cell*, **48**, 422–33.
349. Wasmuth,E. V and Lima,C.D. (2012) Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. *Mol. Cell*, **48**, 133–44.
350. Drazkowska,K., Tomecki,R., Stodus,K., Kowalska,K., Czarnocki-Cieciura,M. and Dziembowski,A. (2013) The RNA exosome complex central channel controls both exonuclease and endonuclease Dis3 activities in vivo and in vitro. *Nucleic Acids Res.*, **41**, 3845–58.
351. Callahan,K.P. and Butler,J.S. (2010) TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. *J. Biol. Chem.*, **285**, 3540–7.
352. Lubas,M., Christensen,M.S., Kristiansen,M.S., Domanski,M., Falkenby,L.G., Lykke-Andersen,S., Andersen,J.S., Dziembowski,A. and Jensen,T.H. (2011) Interaction profiling identifies the human nuclear exosome targeting complex. *Mol. Cell*, **43**, 624–37.
353. Costello,J.L., Stead,J.A., Feigenbutz,M., Jones,R.M. and Mitchell,P. (2011) The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3’-maturation. *J. Biol. Chem.*, **286**, 4535–43.
354. Ciais,D., Bohnsack,M.T. and Tollervey,D. (2008) The mRNA encoding the yeast ARE-binding protein Cth2 is generated by a novel 3’ processing pathway. *Nucleic Acids Res.*, **36**, 3075–84.

355. Roth,K.M., Byam,J., Fang,F. and Butler,J.S. (2009) Regulation of NAB2 mRNA 3'-end formation requires the core exosome and the Trf4p component of the TRAMP complex. *RNA*, **15**, 1045–58.
356. Allmang,C., Mitchell,P., Petfalski,E. and Tollervey,D. (2000) Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res.*, **28**, 1684–91.
357. Bousquet-Antonelli,C., Presutti,C. and Tollervey,D. (2000) Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell*, **102**, 765–75.
358. Libri,D., Dower,K., Boulay,J., Thomsen,R., Rosbash,M. and Jensen,T.H. (2002) Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol. Cell. Biol.*, **22**, 8254–66.
359. Torchet,C., Bousquet-Antonelli,C., Milligan,L., Thompson,E., Kufel,J. and Tollervey,D. (2002) Processing of 3'-extended read-through transcripts by the exosome can generate functional mRNAs. *Mol. Cell*, **9**, 1285–96.
360. Szczepinska,T., Kalisiak,K., Tomecki,R., Labno,A., Borowski,L.S., Kulinski,T., Adamska,D., Kosinska,J. and Dziembowski,A. (2015) DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts. *Genome Res.*, **25**, 1622–33.
361. Callahan,K.P. and Butler,J.S. (2008) Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. *Nucleic Acids Res.*, **36**, 6645–55.
362. Basu,U., Meng,F.L., Keim,C., Grinstein,V., Pefanis,E., Eccleston,J., Zhang,T., Myers,D., Wasserman,C.R., Wesemann,D.R., *et al.* (2011) The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell*, **144**, 353–63.
363. Reimer,G. (1990) Autoantibodies against nuclear, nucleolar, and mitochondrial antigens in systemic sclerosis (scleroderma). *Rheum. Dis. Clin. North Am.*, **16**, 169–83.
364. Chapman,M.A., Lawrence,M.S., Keats,J.J., Cibulskis,K., Sougnez,C., Schinzel,A.C., Harview,C.L., Brunet,J.P., Ahmann,G.J., Adli,M., *et al.* (2011) Initial genome sequencing and analysis of multiple myeloma. *Nature*, **471**, 467–72.
365. Schmidt,K. and Butler,J.S. (2013) Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity. *Wiley Interdiscip. Rev. RNA*, **4**, 217–31.
366. Vanácová,S., Wolf,J., Martin,G., Blank,D., Dettwiler,S., Friedlein,A., Langen,H., Keith,G. and Keller,W. (2005) A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol.*, **3**, e189.
367. Holub,P., Lalakova,J., Cerna,H., Pasulka,J., Sarazova,M., Hrazdilova,K., Arce,M.S., Hobor,F., Stefl,R. and Vanacova,S. (2012) Air2p is critical for the assembly and RNA-binding of the TRAMP complex and the KOW domain of Mtr4p is crucial for exosome activation. *Nucleic Acids Res.*, **40**, 5679–93.
368. Jia,H., Wang,X., Liu,F., Guenther,U.P., Srinivasan,S., Anderson,J.T. and Jankowsky,E. (2011) The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex. *Cell*, **145**, 890–901.
369. Mangus,D.A., Evans,M.C. and Jacobson,A. (2003) Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.*, **4**, 223.
370. Hamill,S., Wolin,S.L. and Reinisch,K.M. (2010) Structure and function of the polymerase core of TRAMP, a RNA surveillance complex. *Proc. Natl. Acad. Sci.*, **107**, 15045–50.
371. Fasken,M.B., Leung,S.W., Banerjee,A., Kodani,M.O., Chavez,R., Bowman,E.A., Purohit,M.K., Rubinson,M.E., Rubinson,E.H. and Corbett,A.H. (2011) Air1 zinc knuckles 4 and 5 and a conserved IWRXY motif are critical for the function and integrity of the Trf4/5-Air1/2-Mtr4 polyadenylation (TRAMP) RNA quality control complex. *J. Biol. Chem.*, **286**, 37429–45.
372. LaCava,J., Houseley,J., Saveanu,C., Petfalski,E., Thompson,E., Jacquier,A. and Tollervey,D. (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell*, **121**, 713–24.
373. Houseley,J. and Tollervey,D. (2006) Yeast Trf5p is a nuclear poly(A) polymerase. *EMBO Rep.*, **7**, 205–11.
374. Schmidt,K., Xu,Z., Mathews,D.H. and Butler,J.S. (2012) Air proteins control differential

TRAMP substrate specificity for nuclear RNA surveillance. *RNA*, **18**, 1934–45.

375. Houseley,J., Kotovic,K., El Hage,A. and Tollervey,D. (2007) Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. *EMBO J.*, **26**, 4996–5006.
376. Egecioglu,D.E., Henras,A.K. and Chanfreau,G.F. (2006) Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. *RNA*, **12**, 26–32.
377. Arigo,J.T., Eyler,D.E., Carroll,K.L. and Corden,J.L. (2006) Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol. Cell*, **23**, 841–51.
378. Walowsky,C., Fitzhugh,D.J., Castaño,I.B., Ju,J.Y., Levin,N.A. and Christman,M.F. (1999) The topoisomerase-related function gene TRF4 affects cellular sensitivity to the antitumor agent camptothecin. *J. Biol. Chem.*, **274**, 7302–8.
379. Dez,C., Houseley,J. and Tollervey,D. (2006) Surveillance of nuclear-restricted pre-ribosomes within a subnucleolar region of Saccharomyces cerevisiae. *EMBO J.*, **25**, 1534–46.
380. Weir,J.R., Bonneau,F., Hentschel,J. and Conti,E. (2010) Structural analysis reveals the characteristic features of Mtr4, a DExH helicase involved in nuclear RNA processing and surveillance. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 12139–44.
381. Jackson,R.N., Klauer,A.A., Hintze,B.J., Robinson,H., van Hoof,A. and Johnson,S.J. (2010) The crystal structure of Mtr4 reveals a novel arch domain required for rRNA processing. *EMBO J.*, **29**, 2205–16.
382. de la Cruz,J., Kressler,D., Tollervey,D. and Linder,P. (1998) Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. *EMBO J.*, **17**, 1128–40.
383. Kadaba,S., Wang,X. and Anderson,J.T. (2006) Nuclear RNA surveillance in Saccharomyces cerevisiae: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. *RNA*, **12**, 508–21.
384. San Paolo,S., Vanacova,S., Schenk,L., Scherrer,T., Blank,D., Keller,W. and Gerber,A.P. (2009) Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. *PLoS Genet.*, **5**, e1000555.
385. Wang,X., Jia,H., Jankowsky,E. and Anderson,J.T. (2008) Degradation of hypomodified tRNA(iMet) in vivo involves RNA-dependent ATPase activity of the DExH helicase Mtr4p. *RNA*, **14**, 107–16.
386. Jia,H., Wang,X., Anderson,J.T. and Jankowsky,E. (2012) RNA unwinding by the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 7292–7.
387. Schmid,M., Poulsen,M.B., Olszewski,P., Pelechano,V., Saguez,C., Gupta,I., Steinmetz,L.M., Moore,C. and Jensen,T.H. (2012) Rrp6p controls mRNA poly(A) tail length and its decoration with poly(A) binding proteins. *Mol. Cell*, **47**, 267–80.
388. Inoue,K., Mizuno,T., Wada,K. and Hagiwara,M. (2000) Novel RING finger proteins, Air1p and Air2p, interact with Hmt1p and inhibit the arginine methylation of Npl3p. *J. Biol. Chem.*, **275**, 32793–9.
389. Vasiljeva,L., Kim,M., Terzi,N., Soares,L.M. and Buratowski,S. (2008) Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. *Mol. Cell*, **29**, 313–23.
390. Shcherbik,N., Wang,M., Lapik,Y.R., Srivastava,L. and Pestov,D.G. (2010) Polyadenylation and degradation of incomplete RNA polymerase I transcripts in mammalian cells. *EMBO Rep.*, **11**, 106–11.
391. Rammelt,C., Bilen,B., Zavolan,M. and Keller,W. (2011) PAPD5, a noncanonical poly(A) polymerase with an unusual RNA-binding motif. *RNA*, **17**, 1737–46.
392. Nguyen,D., Grenier St-Sauveur,V., Bergeron,D., Dupuis-Sandoval,F., Scott,M.S. and Bachand,F. (2015) A polyadenylation-dependent 3' end maturation pathway is required for the synthesis of the human telomerase RNA. *Cell Rep.*, doi: 10.1016/j.celrep.2015.11.003.
393. Watkins,N.J. and Bohnsack,M.T. (2012) The box C/D and H/ACA snoRNPs: Key players in the

modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA*, **3**, 397–414.

394. Galej,W.P., Nguyen,T.H.D., Newman,A.J. and Nagai,K. (2014) Structural studies of the spliceosome: zooming into the heart of the machine. *Curr. Opin. Struct. Biol.*, **25**, 57–66.
395. Nguyen,T.H.D., Galej,W.P., Bai,X., Savva,C.G., Newman,A.J., Scheres,S.H.W. and Nagai,K. (2015) The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature*, **523**, 47–52.
396. Montemayor,E.J., Curran,E.C., Liao,H.H., Andrews,K.L., Treba,C.N., Butcher,S.E. and Brow,D.A. (2014) Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. *Nat. Struct. Mol. Biol.*, **21**, 544–51.
397. Guthrie,C. and Patterson,B. (1988) Spliceosomal snRNAs. *Annu. Rev. Genet.*, **22**, 387–419.
398. Neuenkirchen,N., Chari,A. and Fischer,U. (2008) Deciphering the assembly pathway of Sm-class U snRNPs. *FEBS Lett.*, **582**, 1997–2003.
399. Fica,S.M., Tuttle,N., Novak,T., Li,N.-S., Lu,J., Koodathingal,P., Dai,Q., Staley,J.P. and Piccirilli,J.A. (2013) RNA catalyses nuclear pre-mRNA splicing. *Nature*, **503**, 229–34.
400. Rashid,R., Liang,B., Baker,D.L., Youssef,O.A., He,Y., Phipps,K., Terns,R.M., Terns,M.P. and Li,H. (2006) Crystal structure of a Cbf5-Nop10-Gar1 complex and implications in RNA-guided pseudouridylation and dyskeratosis congenita. *Mol. Cell*, **21**, 249–60.
401. Li,L. and Ye,K. (2006) Crystal structure of an H/ACA box ribonucleoprotein particle. *Nature*, **443**, 302–7.
402. Reichow,S.L., Hamma,T., Ferré-D'Amaré,A.R. and Varani,G. (2007) The structure and function of small nucleolar ribonucleoproteins. *Nucleic Acids Res.*, **35**, 1452–64.
403. van Nues,R.W., Granneman,S., Kudla,G., Sloan,K.E., Chicken,M., Tollervey,D. and Watkins,N.J. (2011) Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *EMBO J.*, **30**, 2420–30.
404. Watkins,N.J., Ségault,V., Charpentier,B., Nottrott,S., Fabrizio,P., Bachi,A., Wilm,M., Rosbash,M., Branlant,C. and Lührmann,R. (2000) A common core RNP structure shared between the small nucleoar box C/D RNPs and the spliceosomal U4 snRNP. *Cell*, **103**, 457–66.
405. Aittaleb,M., Rashid,R., Chen,Q., Palmer,J.R., Daniels,C.J. and Li,H. (2003) Structure and function of archaeal box C/D sRNP core proteins. *Nat. Struct. Biol.*, **10**, 256–63.
406. Granneman,S., Kudla,G., Petfalski,E. and Tollervey,D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 9613–8.
407. Lübben,B., Marshallsay,C., Rottmann,N. and Lührmann,R. (1993) Isolation of U3 snoRNP from CHO cells: a novel 55 kDa protein binds to the central part of U3 snoRNA. *Nucleic Acids Res.*, **21**, 5377–85.
408. Granneman,S., Pruijn,G.J.M., Horstman,W., van Venrooij,W.J., Luhrmann,R. and Watkins,N.J. (2002) The hU3-55K protein requires 15.5K binding to the box B/C motif as well as flanking RNA elements for its association with the U3 small nucleolar RNA in Vitro. *J. Biol. Chem.*, **277**, 48490–500.
409. Granneman,S., Vogelzangs,J., Lührmann,R., van Venrooij,W.J., Pruijn,G.J.M. and Watkins,N.J. (2004) Role of pre-rRNA base pairing and 80S complex formation in subnucleolar localization of the U3 snoRNP. *Mol. Cell. Biol.*, **24**, 8600–10.
410. Cléry,A., Senty-Ségault,V., Leclerc,F., Raué,H.A. and Branlant,C. (2007) Analysis of sequence and structural features that identify the B/C motif of U3 small nucleolar RNA as the recognition site for the Snu13p-Rrp9p protein pair. *Mol. Cell. Biol.*, **27**, 1191–206.
411. Phipps,K.R., Charette,J.M. and Baserga,S.J. The small subunit processome in ribosome biogenesis—progress and prospects. *Wiley Interdiscip. Rev. RNA*, **2**, 1–21.
412. Knox,A.A., McKeegan,K.S., Debieux,C.M., Traynor,A., Richardson,H. and Watkins,N.J. (2011) A weak C' box renders U3 snoRNA levels dependent on hU3-55K binding. *Mol. Cell. Biol.*, **31**, 2404–12.
413. Lafontaine,D.L., Bousquet-Antonelli,C., Henry,Y., Caizergues-Ferrer,M. and Tollervey,D.

(1998) The box H + ACA snoRNAs carry Cbf5p, the putative rRNA pseudouridine synthase. *Genes Dev.*, **12**, 527–37.

414. Zebarjadian,Y., King,T., Fournier,M.J., Clarke,L. and Carbon,J. (1999) Point mutations in yeast CBF5 can abolish in vivo pseudouridylation of rRNA. *Mol. Cell. Biol.*, **19**, 7461–72.
415. Tollervey,D., Lehtonen,H., Jansen,R., Kern,H. and Hurt,E.C. (1993) Temperature-sensitive mutations demonstrate roles for yeast fibrillarin in pre-rRNA processing, pre-rRNA methylation, and ribosome assembly. *Cell*, **72**, 443–57.
416. Kiss-László,Z., Henry,Y., Bachellerie,J.P., Caizergues-Ferrer,M. and Kiss,T. (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell*, **85**, 1077–88.
417. Galardi,S., Fatica,A., Bachi,A., Scaloni,A., Presutti,C. and Bozzoni,I. (2002) Purified box C/D snoRNPs are able to reproduce site-specific 2'-O-methylation of target RNA in vitro. *Mol. Cell. Biol.*, **22**, 6663–8.
418. Henras,A.K., Soudet,J., Gérus,M., Lebaron,S., Caizergues-Ferrer,M., Mougin,A. and Henry,Y. (2008) The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell. Mol. Life Sci.*, **65**, 2334–59.
419. Tycowski,K.T., You,Z.H., Graham,P.J. and Steitz,J.A. (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs. *Mol. Cell*, **2**, 629–38.
420. Collins,K. (2006) The biogenesis and regulation of telomerase holoenzymes. *Nat. Rev. Mol. Cell Biol.*, **7**, 484–94.
421. Esakova,O. and Krasilnikov,A.S. (2010) Of proteins and RNA: the RNase P/MRP family. *RNA*, **16**, 1725–47.
422. Dieci,G., Preti,M. and Montanini,B. (2009) Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics*, **94**, 83–8.
423. Richard,P. and Kiss,T. (2006) Integrating snoRNP assembly with mRNA biogenesis. *EMBO Rep.*, **7**, 590–92.
424. Preti,M., Guffanti,E., Valitutto,E. and Dieci,G. (2006) Assembly into snoRNP controls 5'-end maturation of a box C/D snoRNA in Saccharomyces cerevisiae. *Biochem. Biophys. Res. Commun.*, **351**, 468–73.
425. Preti,M., Ribeyre,C., Pascali,C., Bosio,M.C., Cortelazzi,B., Rougemont,J., Guarnera,E., Naef,F., Shore,D. and Dieci,G. (2010) The Telomere-Binding Protein Tbf1 Demarcates snoRNA Gene Promoters in Saccharomyces cerevisiae. *Mol. Cell*, **38**, 614–20.
426. Yang,P.K., Hoareau,C., Froment,C., Henry,Y., Chanfreau,G. and Monsarrat,B. (2005) Cotranscriptional recruitment of the pseudouridylsynthetase Cbf5p and of the RNA binding protein Naf1p during H/ACA snoRNP assembly. *Mol. Cell. Biol.*, **25**, 3295–304.
427. Houalla,R., Devaux,F., Fatica,A., Kufel,J., Barrass,D., Torchet,C. and Tollervey,D. (2006) Microarray detection of novel nuclear RNA substrates for the exosome. *Yeast*, **23**, 439–54.
428. Larochelle,M., Lemay,J.F. and Bachand,F. (2012) The THO complex cooperates with the nuclear RNA surveillance machinery to control small nucleolar RNA expression. *Nucleic Acids Res.*, **40**, 10240–53.
429. Kufel,J., Allmang,C., Chanfreau,G., Petfalski,E., Lafontaine,D.L. and Tollervey,D. (2000) Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol. Cell. Biol.*, **20**, 5415–24.
430. Xue,D., Rubinson,D.A., Pannone,B.K., Yoo,C.J. and Wolin,S.L. (2000) U snRNP assembly in yeast involves the La protein. *EMBO J.*, **19**, 1650–60.
431. van Hoof,A., Lennertz,P. and Parker,R. (2000) Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. *EMBO J.*, **19**, 1357–65.
432. Machyna,M., Heyn,P. and Neugebauer,K.M. (2013) Cajal bodies: where form meets function. *Wiley Interdiscip. Rev. RNA*, **4**, 17–34.
433. Yang,P.K., Rotondo,G., Porras,T., Legrain,P. and Chanfreau,G. (2002) The Shq1p.Naf1p

complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J. Biol. Chem.*, **277**, 45235–42.

434. Fatica,A., Dlakić,M. and Tollervey,D. (2002) Naf1 p is a box H/ACA snoRNP assembly factor. *RNA*, **8**, 1502–14.
435. Lindstrom,D.L., Squazzo,S.L., Muster,N., Burckin,T.A., Wachter,K.C., Emigh,C.A., McCleery,J.A., Yates,J.R. and Hartzog,G.A. (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. *Mol. Cell. Biol.*, **23**, 1368–78.
436. Bratkovič,T. and Rogelj,B. (2011) Biology and applications of small nucleolar RNAs. *Cell. Mol. Life Sci.*, **68**, 3843–51.
437. Kiss,T., Fayet-Lebaron,E. and Jády,B.E. (2010) Box H/ACA small ribonucleoproteins. *Mol. Cell*, **37**, 597–606.
438. Rothé,B., Saliou,J.M., Quinternet,M., Back,R., Tiotiu,D., Jacquemin,C., Loegler,C., Schlotter,F., Peña,V., Eckert,K., *et al.* (2014) Protein Hit1, a novel box C/D snoRNP assembly factor, controls cellular concentration of the scaffolding protein Rsa1 by direct interaction. *Nucleic Acids Res.*, **42**, 10731–47.
439. Prieto,M.B., Georg,R.C., Gonzales-Zubiate,F.A., Luz,J.S. and Oliveira,C.C. (2015) Nop17 is a key R2TP factor for the assembly and maturation of box C/D snoRNP complex. *BMC Mol. Biol.*, **16**, 7.
440. von Morgen,P., Hořejší,Z. and Macurek,L. (2015) Substrate recognition and function of the R2TP complex in response to cellular stress. *Front. Genet.*, **6**, 69.
441. Hirose,T. and Steitz,J.A. (2001) Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 12914–9.
442. Hirose,T., Shu,M.D. and Steitz,J.A. (2003) Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. *Mol. Cell*, **12**, 113–23.
443. Eggert,C., Chari,A., Laggerbauer,B. and Fischer,U. (2006) Spinal muscular atrophy: The RNP connection. *Trends Mol. Med.*, **12**, 113–21.
444. Li,D.K., Tisdale,S., Lotti,F. and Pellizzoni,L. (2014) SMN control of RNP assembly: from post-transcriptional gene regulation to motor neuron disease. *Semin. Cell Dev. Biol.*, **32**, 22–9.
445. Bizarro,J., Dodré,M., Huttin,A., Charpentier,B., Schlotter,F., Branlant,C., Verheggen,C., Massenet,S. and Bertrand,E. (2015) NUFIP and the HSP90/R2TP chaperone bind the SMN complex and facilitate assembly of U4-specific proteins. *Nucleic Acids Res.*, **43**, 8973–89.
446. Szczepaniak,S.A., Zuberek,J., Darzynkiewicz,E., Kufel,J. and Jemielity,J. (2012) Affinity resins containing enzymatically resistant mRNA cap analogs - a new tool for the analysis of cap-binding proteins. *RNA*, **18**, 1421–32.
447. Webb,N.R., Chari,R. V, DePillis,G., Kozarich,J.W. and Rhoads,R.E. (1984) Purification of the messenger RNA cap-binding protein using a new affinity medium. *Biochemistry*, **23**, 177–81.
448. Aguilar,R.M., Bustamante,J.J., Hernandez,P.G., Martinez,A.O. and Haro,L.S. (1999) Precipitation of dilute chromatographic samples (ng/ml) containing interfering substances for SDS-PAGE. *Anal. Biochem.*, **267**, 344–50.
449. Szczepaniak,S.A., Jemielity,J., Zuberek,J., Kufel,J. and Darzynkiewicz,E. (2008) Bisphosphonate mRNA cap analog attached to Sepharose for affinity chromatography of decapping enzymes. *Nucleic Acids Symp. Ser. (Oxf).*, **52**, 295-6.
450. Gelperin,D.M., White,M.A., Wilkinson,M.L., Kon,Y., Kung,L.A., Wise,K.J., Lopez-Hoyo,N., Jiang,L., Piccirillo,S., Yu,H., *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev.*, **19**, 2816–26.
451. Kiraga-Motoszko,K., Stepinski,J., Niedzwiecka,A., Jemielity,J., Wszelaka-Rylik,M., Stolarski,R., Zielenkiewicz,W. and Darzynkiewicz,E. (2003) Interaction between yeast eukaryotic initiation factor eIF4E and mRNA 5' cap analogues differs from that for murine eIF4E. *Nucleosides. Nucleotides Nucleic Acids*, **22**, 1711–4.
452. Niedzwiecka,A., Marcotrigiano,J., Stepinski,J., Jankowska-Anyszka,M., Wyslouch-Cieszynska,A., Dadlez,M., Gingras,A.C., Mak,P., Darzynkiewicz,E., Sonenberg,N., *et al.* (2002)

Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J. Mol. Biol.*, **319**, 615–35.

453. Mitchell,P., Petfalski,E., Houalla,R., Podtelejnikov,A., Mann,M. and Tollervey,D. (2003) Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. *Mol. Cell. Biol.*, **23**, 6982–92.
454. Churchman,L.S. and Weissman,J.S. (2011) Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature*, **469**, 368–73.
455. Rasmussen,T.P. and Culbertson,M.R. (1998) The putative nucleic acid helicase Sen1p is required for formation and stability of termini and for maximal rates of synthesis and levels of accumulation of small nucleolar RNAs in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, **18**, 6885–96.
456. Baudin-Baillieu,A., Tollervey,D., Cullin,C. and Lacroute,F. (1997) Functional analysis of Rrp7p, an essential yeast protein involved in pre-rRNA processing and ribosome assembly. *Mol. Cell. Biol.*, **17**, 5023–32.
457. Parker,J.L., Bielen,A.B., Dikic,I. and Ulrich,H.D. (2007) Contributions of ubiquitin- and PCNA-binding domains to the activity of Polymerase eta in Saccharomyces cerevisiae. *Nucleic Acids Res.*, **35**, 881–9.
458. Mumberg,D., Müller,R. and Funk,M. (1994) Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res.*, **22**, 5767–8.
459. Van Driessche,B., Tafforeau,L., Hentges,P., Carr,A.M. and Vandenhaute,J. (2005) Additional vectors for PCR-based gene tagging in Saccharomyces cerevisiae and Schizosaccharomyces pombe using nourseothricin resistance. *Yeast*, **22**, 1061–8.
460. Sambrook,J. and Russell,D.W. (2001) Molecular Cloning: A Laboratory Manual, Volume 1 CSHL Press.
461. Longtine,M.S., McKenzie,A., Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P. and Pringle,J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast*, **14**, 953–61.
462. Gietz,D., St Jean,A., Woods,R.A. and Schiestl,R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.*, **20**, 1425.
463. Schmitt,M.E., Brown,T.A. and Trumpower,B.L. (1990) A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae. *Nucleic Acids Res.*, **18**, 3091–2.
464. Tollervey,D. (1987) A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. *EMBO J.*, **6**, 4169–75.
465. Marcotrigiano,J., Gingras,A.C., Sonenberg,N. and Burley,S.K. (1997) Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell*, **89**, 951–61.
466. Cohen,L.S., Mikhli,C., Friedman,C., Jankowska-Anyszka,M., Stepinski,J., Darzynkiewicz,E. and Davis,R.E. (2004) Nematode m7GpppG and m3(2,2,7)GpppG decapping: activities in Ascaris embryos and characterization of C. elegans scavenger DcpS. *RNA*, **10**, 1609–24.